UNIVERSITY OF MINNESOTA RECEIVED JUL 11 1942 LIBRARY

The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Volume X | June, 1942 | Number 4

CURRICULUM CONSTRUCTION

CONTENTS

PAGE

A Study Relating to the Change in the Newspaper Reading Interests of Secondary School Students Since the Entrance of the United States into World War II: *Homer F. Tangney* ---------- 195

Some Effects of a Five-Year Developmental Experiment Sponsored by a Private Social Agency in a Public School: *Jesse Batley Rhinehart* ---------- 200

A Frame of Reference for the Organization and Administration of Character Education in the Elementary School: *Isidore Kayfetz* ---------- 206

Objectivity in the Evaluation of Personality: *T. W. Richards and Willis Ellington* 228

A Study of Children's Hearing as It Relates to Reading: *Helen Kennedy* ---------- 238

Factors Affecting Efficiency of Inductive Reasoning: *Louis Long and Livingston Welch* ---------- 252

Factors Contributing to Achievement in the Study of First Semester College German: *R. P. Larsen, J. R. Wittenborn, and E. G. Giesecke* ---------- 265

$5.00 a Year | Published Quarterly | $1.50 a Copy

Edited and Published by A. S. Barr, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Entered as second-class matter October 17, 1935 at the post office at Madison, Wisconsin under the act of March 3, 1879.

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison, Wis.

Carter V. Good, Professor of Education, University of Cincinnati, Cincinnati, Ohio. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Assistant Director, Bureau of Reference, Research and Statistics, Board of Education of the City of New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.

Edward E. Cureton, Senior Educational Statistician, U. S. Office of Education, Washington, D. C. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

CONTRIBUTING EDITORS

Gilbert L. Betts, Supervisor of Graduate Research in Education, Colorado State College, Fort Collins, Colorado.

William A. Brownell, Professor of Educational Psychology, Duke University, Durham, North Carolina.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Otis W. Caldwell, General Secretary, The American Association for the Advancement of Science, Boyce Thompson Institute for Plant Research, Yonkers, New York.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, Associate Professor of Psychology, College of Agriculture, and Research Associate, Institute of Child Welfare, University of California, Berkeley, California.

Stephen M. Corey, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Robert A. Davis, Professor of Education, Director of Bureau of Educational Research, University of Colorado, Boulder, Colorado.

Harl R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Jack W. Dunlap, Associate Professor of Educational Psychology, University of Rochester, Rochester, New York.

Harold A. Edgerton, Assistant Professor of Psychology, Ohio State University, Columbus, Ohio.

Alvin C. Eurich, Professor of Education, Stanford University, Stanford University, California.

John C. Flanagan, Assistant Chief, Research Section, Medical Division, Office, Chief of the Air Corps, War Department, Washington, D. C.

Kai Jensen, Associate Professor of Education, University of Wisconsin, Madison, Wisconsin.

Harold E. Jones, Professor of Psychology, Director of Research, Institute of Child Welfare, University of California, Berkeley, California.

Noel Keys, Professor of Education, University of California, Berkeley, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

T. E. Newland, Pennsylvania State Department of Education, Chief of Special Education, Harrisburg, Pennsylvania.

C. W. Odell, Associate Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Professor of Education, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

W. E. Peik, Dean and Professor of Education, University of Minnesota, Minneapolis, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

Clarence E. Ragsdale, Associate Professor of Education, University of Wisconsin, Madison, Wisconsin.

H. H. Remmers, Director, Division of Educational Reference, Professor of Education and Psychology, Purdue University, Lafayette, Indiana.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

G. M. Ruch, Chief of Research and Statistical Service, U. S. Office of Education, Washington, D. C.

P. J. Rulon, Assistant Professor of Education, Graduate School of Education, Harvard University, Cambridge, Massachusetts.

Douglas E. Scates, Associate Professor of Education, Duke University, Durham, North Carolina.

David Segel, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Clinic of Child Development, Research Associate, Yale University, New Haven, Connecticut.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

T. L. Torgerson, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Helen M. Walker, Professor of Education, Teachers College, Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Professor of Education, Boston University, Boston, Massachusetts.

Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

DEMOCRAT PRINTING COMPANY
MADISON, WISCONSIN

Journal of Experimental Education

Volume X · JUNE, 1942 · Number 4

A STUDY RELATING TO THE CHANGE IN THE NEWSPAPER READING INTERESTS OF SECONDARY SCHOOL STUDENTS SINCE THE ENTRANCE OF THE UNITED STATES INTO WORLD WAR II

HOMER F. TANGNEY
West Milwaukee High School
Milwaukee, Wis.

Since "Pearl Harbor", December 7, 1941, Americans have had a decidedly more vital interest in the day-by-day happenings of the world than before. Secondary school students constitute an important part of the newspaper reading public, and any determination of a change of their reading interests should indicate a trend in the changing interests of the general public. The purpose of this study is to determine whether or not there has been a change in the newspaper reading interests of secondary school students, and, if so, the extent of the changing interests and the trends in these changes of interests.

THE MEASURING INSTRUMENT

The measuring instrument used in this study was an interest questionnaire, comprising fifty items of features of the daily newspapers. The items are those listed in Tables I and II. Where it was considered advisable, the items were followed by brief explanations, such as, "Radio: programs, reviews, news of stars, etc." The questionnaire was of the L I D type, and the students were asked to indicate whether they "liked", were "indifferent to", or "disliked" each item by checking the proper column. The tabulations for this study were based entirely upon the "Like", responses. Inasmuch as this was a comparative study of interest, no attempt was made to evaluate the "Indifferent" and "Dislike" responses. It might be added that all figures in Tables I and II have been reduced to percentages.

PROCEDURE

The comparison made in this study is based upon two administrations of the questionnaire. The questionnaire was first administered early in May, 1941, and was given to about 275 students. At that time it was given for a purpose other than this study. Following the entrance of the United States in World War II it was decided to administer the questionnaire a second time. Six weeks were allowed to elapse from our entrance into the war before the questionnaire was given the second time, so that the first burst of enthusiastic interest might settle down to a normal wartime interest. The questionnaire was given late in January, 1942, to 260 students. More than eight months' time elapsed between the two administrations. No attempt was made to give the questionnaire to the same students each time. In fact, that would have been impossible as the seniors who filled out the questionnaire on the first administration had been graduated and the freshmen on the second administration were not in school when the first administration of the questionnaire was made. Many of those who took the first administration as freshmen took it the second time as sophomores. Many sophomores at the first administration were juniors at the second, and many juniors at the first were seniors at the second administration. A probable estimate is that about one-third of students filled out the questionnaire both times.

The questionnaires of 200 students from each administration of the questionnaire were used to make the comparison. Twenty-five boys and twenty-five girls were selected at random for each of the four years of high school. The balance of the questionnaires was discarded. The number twenty-five was fixed as the sampling from each group for

three reasons: (1) that number is generally conceded to give reliable results in a study of this kind, (2) in two groups only that number were available, (3) the use of twenty-five cases made for easy computation.

The tabulations of the results of these two administrations of the questionnaire are shown in Table I and II, with all figures reduced to percentages.

Table I shows the scores (in percentages) for girls and boys, by grades, and for both administrations of the questionnaire. It also shows the totals for each administration, and the percentage increase or decrease in reading interest for each of the fifty items of the questionnaire.

Table II shows the comparative percentages on the two administrations of the questionnaire by grades, and separately for girls and boys. It also shows the percentage of increase or decrease in reading interest for totals for both girls and boys for each of the items of the questionnaire.

Findings

The total tabulation shows an increased interest in fifteen of the fifty items, ranging from .5% to 16.5% (average 3.97%), and decrease in thirty-two items, with the same range (average 3.92%). Four items—illustrations, radio, war news, and headlines—showed an increase of more than 10%, and two items—school news and fraternal news—showed a decrease of more than 10% in interest. These changes are large enough to be considered as definite indications of change in reading interest. Eight items—sports, travel news, college news, crime news, scandal, market quotations, employment and labor news—showed decreases of interest of between 5% and 10%, certainly significant indications of decreased interest. To these should be added increased interest in foreign news and religion and decreases in the theater and medical articles which were slightly under the 5% level, but are also significant. Three items showed no change, and the change in fourteen items was 1% or less. These changes are too small to be significant and could be expected to appear in any sampling. The remainder of the changes show increases or decreases in interest ranging from 1% to 4%, and may or may not be significant.

The separate tabulations for girls and boys show a decided difference in sex interests and the change of interests. Girls showed an increased interest in twenty items and a decreased interest in twenty-five items, while boys showed increased interest in eighteen items and decreased interest in twenty-nine items. The average increase for girls was 6.5% and 4.4% for boys, while the average decrease in interest was 5.3% and 5.1 for boys. The average change of both increase and decrease was 5.3% for girls and 5.1% for boys. Girls showed a much more marked increased interest in illustrations, war news, foreign news, and headlines than did the boys, and a decidedly decreased interest in such items as fraternal news, medical, scientific articles, letters, want ads, and commercial advertisements. Boys showed a marked increased interest in local news, letters, and want ads, and a more marked decreased interest in such items as the theater, accident reports, crime news, scandal, headlines, speeches, employment, labor news, commodity market, and medical articles.

In general it may be said that the findings of the study bears out the findings of Fendrick,[1] namely: (1) that certain items are more popular than others, (2) that there are progressive age levels of reading interests, and (3) that there is a decided sex difference in the reading interests of secondary school students.

A perfect correlation of the items would be impossible because Fendrick's study used thirty-five items and the present study used fifty items. However, the first three items in popularity were the same in both studies, namely, photographs, comics, and cartoons, and in the same order. There is a strong similarity of the popularity of the corresponding items in the two studies. The age levels of reading interest and the advance in interest from grade level to grade level is shown in a number of the items, such as radio and local news. Girls showed an increase in interest as the grade level advances in such items as society news, styles and fashions, household hints, and beauty advice. Seniors showed a markedly increased interest in employment and want ads. Sex difference was clearly shown. Girls showed a more marked interest than boys in movies, theater, musicals, radio,

[1] Paul Fendrick. "Newspaper Reading Interests of High School and College Students," *Journal of Educational Research*, XXXIV (March, 1941), 522–530.

TABLE I

SHOWING THE RAW SCORES, THE TOTAL SCORES AND THE

TABLE I

SHOWING THE RAW SCORES, THE TOTAL SCORES, AND THE ITEM INCREASE OR DECREASE FOR TWO ADMINISTRATIONS OF A NEWSPAPER INTEREST QUESTIONNAIRE TO 200 SECONDARY SCHOOL STUDENTS

(All Figures Have Been Reduced to Percentages)

Note: (1) First Administration—May, 1941; (2) Second Administration—January, 1942

	9th Grade				10th Grade				11th Grade				12th Grade				Total		
	Girls		Boys		Girls		Boys		Girls		Boys		Girls		Boys		Girls + Boys		
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	
Photographs	88	92	84	88	92	100	100	96	100	100	96	96	96	96	92	88	93.5	94.5	+ 1
Comics	92	96	96	96	96	92	100	92	100	88	100	100	76	88	96	96	94.5	93.5	— 1
Cartoons	80	76	88	92	76	72	96	92	88	84	92	88	76	92	96	84	86.5	85	— 1.5
Illustrations	40	80	72	56	40	52	24	60	40	64	56	64	52	36	68	64	48.5	59.5	+11
Sports	72	76	84	80	68	52	96	88	76	76	84	92	56	40	92	76	78.5	72.5	— 6
Movies	84	88	72	28	88	92	40	64	88	88	68	72	80	80	64	56	70.5	71	— .5
Theater	72	84	52	24	84	84	40	40	80	80	60	72	80	72	56	36	65.5	61.5	— 4
Musical Events	24	16	8	16	32	28	12	12	60	48	32	44	28	48	32	32	26.5	30.5	+ 2
Radio	68	88	64	56	92	92	32	64	72	88	72	84	80	92	80	84	70	81	+11
War News	16	32	60	68	4	36	60	72	16	56	68	72	20	40	68	68	39	55.5	+16.5
Foreign News	24	24	44	52	4	24	36	52	24	36	68	64	24	28	64	48	36.5	41	+ 4.5
Travel News	48	32	48	36	24	28	52	32	44	36	44	56	40	36	48	32	43.5	36	— 7.5
Local News	64	40	44	56	64	68	76	92	72	80	76	80	80	68	84	72	70	69.5	— .5
College News	44	32	40	32	52	52	24	36	64	44	35	44	44	64	60	20	45.5	40.5	— 5
School News	92	80	80	68	92	84	56	76	84	88	76	56	84	80	96	76	82.5	76	— 6.5
Society News	40	36	4	8	48	36	4	8	60	20	12	8	44	48	80	8	37	21.5	—15.5
Fraternal News	8	8	16	12	8	12	8	8	72	8	28	16	80	20	12	16	29	12.5	—16.5
Accident Reports	52	68	60	44	48	64	76	80	60	56	60	56	44	36	52	48	56.5	56.5	
Crime News	40	56	84	48	48	60	76	80	52	56	52	60	48	32	80	40	60	54	— 6
Scandal	36	12	28	12	36	48	44	36	20	36	28	32	28	24	36	16	32	27	— 5
Editorials	48	28	48	24	16	36	68	44	48	48	32	60	36	52	36	36	41.5	41	— .5
Headlines	52	64	72	72	48	88	76	72	64	72	56	72	44	56	68	68	59.5	70.5	+11
Columnists	20	12	36	20	16	16	24	24	24	44	28	40	28	36	44	28	27.5	27.5	
Speeches	8	0	20	12	4	8	8	12	12	20	20	20	16	12	16	12	13	12	— 1
Political Events	16	16	20	36	4	16	24	24	12	8	24	20	8	12	28	12	17	18	— 1
Legislation	12	12	24	20	0	12	24	20	12	12	16	24	12	4	36	24	17	16	— 1
Market Quotations	4	0	12	20	4	0	12	4	8	0	20	8	0	0	36	20	12	6.5	— 5.5
Financial Page	8	4	16	12	0	4	12	8	4	0	16	12	8	4	24	20	11	8	— 3
Industrial News	12	24	52	36	12	8	40	44	16	16	36	36	12	24	48	48	31	32	+ 1
Employment	20	16	40	24	16	16	44	44	24	32	28	40	48	36	68	40	36	31	— 5
Labor News	44	28	48	48	16	32	56	40	20	20	36	36	32	16	60	28	39	31	— 8
Farm News	20	4	28	40	0	4	28	28	0	8	32	20	4	8	40	12	19	15.5	— 3.5
Commodity Market	8	4	24	24	0	0	20	24	4	12	24	20	16	4	16	8	14	12	— 2
Vital Statistics	48	48	28	24	76	72	40	40	80	72	32	48	56	64	36	28	49.5	49.5	
Weather Reports	64	72	60	68	64	80	76	88	84	84	76	68	56	56	80	48	70	70.5	+ .5
Medical	60	16	36	36	52	56	60	60	56	48	36	60	48	56	44	24	49	44.5	— 4.5
Scientific Articles	44	36	56	64	24	32	56	64	24	28	52	52	56	28	52	52	45.5	44.5	— 1
Religion	36	28	28	28	24	44	20	44	32	44	20	20	24	20	28	16	26.5	30.5	+ 4
Book Reviews	48	44	28	20	48	44	8	28	60	64	36	28	52	44	24	28	38	37.5	— .5
Poetry	44	48	24	24	40	28	12	12	52	40	20	12	28	48	32	28	31.5	30	— 1.5
Letters	60	36	44	44	36	60	36	52	60	48	60	40	60	36	28	64	48	47.5	— .5
History-Biography	40	28	56	40	20	40	40	64	32	40	56	52	40	32	56	52	42.5	43.5	+ 1
Humor	92	80	72	84	92	84	92	92	88	92	92	80	80	80	80	84	86	84.5	— 1.5
Art	48	36	32	40	64	52	32	56	52	44	48	24	44	60	44	32	45.5	43	— 2.5
Styles-Fashions	80	68	4	4	96	88	8	24	96	96	28	12	96	96	24	16	54	50.5	— 3.5
Household Hints	56	60	4	8	76	80	0	16	80	56	4	4	76	64	16	8	39	37	— 2
Beauty Advice	56	72	4	12	80	72	4	8	92	80	0	8	88	84	8	12	41.5	43.5	+ 2
Want Ads	40	32	28	32	48	36	36	40	56	44	28	52	52	56	56	56	43	43.5	+ .5
Classified Ads	8	16	24	12	16	24	28	32	36	24	20	32	24	28	32	40	23.5	26	+ 2.5
Commercial Ads	32	28	28	20	76	48	12	32	60	68	28	32	60	48	24	20	40	37	— 3

SHOWING RAW SCORES BY GRADES, AND THE TOTALS, INCREASES AND DECREASES BY SEXES FOR TWO ADMINISTRATIONS OF A NEWSPAPER INTEREST QUESTIONNAIRE TO 200 SECONDARY SCHOOL STUDENTS

(All Figures Have Been Reduced to Percentages)

Note: (1) First Administration—May, 1941; (2) Second Administration—January, 1942

	9 Grade		10 Grade		11 Grade		12 Grade		Girls		Totals	Boys		
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)		(1)	(2)	
Photographs	86	90	96	98	98	98	94	92	94	97	+ 3	93	92	— 1
Comics	94	96	98	92	100	94	86	92	91	91		98	96	— 2
Cartoons	84	84	86	82	90	86	86	88	80	81	+ 1	93	89	— 4
Illustrations	56	68	32	56	48	64	60	50	43	58	+15	54	61	+ 7
Sports	78	78	82	70	80	84	74	58	68	61	— 7	89	84	— 5
Movies	78	58	64	78	78	80	72	68	85	87	+ 2	56	55	— 1
Theater	62	54	62	62	70	76	68	54	79	80	+ 1	52	43	— 9
Musical	16	16	22	20	46	46	30	40	36	35	— 1	21	26	+ 5
Radio	66	72	62	78	72	86	80	88	78	90	+12	62	72	+10
War News	38	50	32	54	42	64	44	54	14	41	+27	64	70	+ 6
Foreign News	34	38	20	38	46	50	44	38	19	28	+ 9	54	54	
Travel News	48	34	38	30	44	46	44	34	39	33	— 6	48	39	— 9
Local News	54	48	70	80	74	80	82	70	70	64	— 6	70	75	+ 5
College News	42	32	38	44	50	44	52	42	51	48	— 3	40	33	— 7
School News	86	74	74	80	80	72	90	78	88	83	— 5	77	69	— 8
Society News	22	22	26	22	36	14	62	28	49	35	—14	25	8	—17
Fraternal News	12	10	8	10	50	12	46	18	42	12	—30	16	13	— 3
Accident Reports	56	56	62	72	60	56	48	42	51	56	+ 5	62	57	— 5
Crime News	62	52	62	70	52	58	64	36	47	51	+ 4	73	57	—16
Scandal	32	12	40	42	24	34	32	20	30	30		34	24	—10
Editorials	48	26	42	40	40	54	36	44	37	41	+ 4	46	41	— 5
Headlines	62	68	62	80	60	72	56	62	52	70	+18	67	71	+ 4
Columnists	28	16	20	20	26	42	36	32	22	27	+ 5	33	28	— 5
Speeches	14	6	6	10	16	20	16	12	10	10		16	14	— 2
Political Events	18	26	14	20	18	14	18	12	10	13	+ 3	24	23	— 1
Legislation	18	16	12	16	14	18	24	14	9	10	+ 1	25	22	— 3
Market Quotations	8	10	8	2	14	4	18	10	4	0	— 4	20	13	— 7
Financial Page	12	8	6	6	10	6	16	12	5	3	— 2	17	13	— 4
Industrial News	32	30	26	26	26	26	30	36	13	18	+ 5	44	41	— 3
Employment	30	20	30	30	26	36	58	38	27	25	— 2	45	37	— 8
Labor News	46	38	36	36	28	28	46	22	28	24	— 4	50	38	—12
Farm News	24	22	14	16	16	14	22	10	6	6		32	25	— 7
Commodity Market	16	14	10	12	14	16	16	6	7	5	— 2	21	19	— 2
Vital Statistics	38	36	58	56	56	60	46	46	65	64	— 1	34	35	+ 1
Weather Reports	62	70	70	84	80	76	68	52	67	73	+ 6	73	68	— 5
Medical	48	26	56	58	46	54	46	40	54	44	—10	44	45	+ 1
Scientific Articles	50	50	40	48	38	40	54	40	37	31	— 6	54	58	+ 4
Religion	32	28	22	44	26	32	26	18	29	34	+ 5	24	27	+ 3
Book Reviews	38	32	28	36	48	46	38	36	52	49	— 3	24	26	+ 2
Poetry	34	36	26	20	36	26	30	38	41	41		22	19	— 3
Letters	52	40	36	56	60	44	44	50	54	45	— 9	42	50	+ 8
History-Biography	48	34	30	52	44	46	48	42	33	35	+ 2	52	52	
Humor	82	82	92	88	90	86	80	82	88	84	— 4	84	85	+ 1
Art	40	38	48	54	50	34	44	46	52	48	— 4	39	38	— 1
Styles-Fashions	44	36	52	56	64	54	60	56	92	87	— 5	16	14	— 2
Household Hints	30	34	38	48	42	30	46	36	72	65	— 7	6	9	+ 3
Beauty Advice	30	42	42	40	46	44	48	48	79	77	— 2	4	10	+ 6
Want Ads	34	32	42	38	42	48	54	56	49	42	— 7	37	45	+ 8
Classified Ads	16	14	22	28	28	28	28	34	21	23	+ 2	26	29	+ 3
Commercial Ads	30	24	44	40	44	50	42	34	57	48	— 9	23	26	+ 3

society news, vital statistics, book reports, poetry, art, styles and fashions, household hints, beauty advice, and commercial advertisements. While boys showed a decidedly larger interest than girls in sports, war news, foreign news, crime news, political news, legislation, market quotations, financial news, industrial news, employment, labor news, farm news, commodity market, scientific articles and history-biography.

Conclusions

From these findings it seems reasonable to draw the following conclusions:

1. That the entrance of the United States into World War II has influenced a change in the newspaper reading interests of secondary school students. This is based on the marked increase in interest shown in such items as illustrations, radio, war news, and headlines and by the marked decrease of interest in such items as society news and farternal news. This conclusion is further strengthened by the lesser, but significant, increased interest shown in foreign news and religion and by the decreased interest in sports, travel news, college news, school news, scandal, market quotations, employment and labor news.

2. The entrance of the United States into the World War II has made a greater change in the newspaper reading interests of girls than boys. This is shown by the fact that the interests of girls increased in a larger number of items, by more marked increases and decreases in individual items, such as increases in war news and headlines and decrease in fraternal news, and the fact that the average change of interest is higher for girls than for boys.

3. On the whole, there were more declines of interest than gains, but this trend was more marked among the boys than among the girls.

4. That the definite and significant changes in newspaper reading interests of secondary school students, as shown by this study, are of such a nature that they can be explained by the entrance of our country into the World War II.

5. This study bears out the findings of Fendrick as to the most popular features of the daily newspapers, that there are progressive age levels of reading interests, and that there is a decided sex difference in the newspaper reading interests of secondary school students.

SOME EFFECTS OF A FIVE-YEAR DEVELOPMENTAL EXPERIMENT SPONSORED BY A PRIVATE SOCIAL AGENCY IN A PUBLIC SCHOOL*

JESSE BATLEY RHINEHART
Hiram House, Cleveland, Ohio

PURPOSE

The purpose of the present study is to present a five-year evaluation of a program of child development conducted by the staff of a private social agency in a public school. Much research has been conducted for the purpose of studying the processes of growth and for the purpose of determining optimum conditions of development. Various standards have thus been evolved, and, in consequence, many service programs have been developed for the purpose of attempting to improve, through educational methods, the conditions under which children develop and grow to maturity. However, comparatively little has been done in the way of determining just how effective such educational programs are, i.e., how much the children served by such educational programs have improved.

The present study, then, is unique in that it attempts to determine the extent to which one program of child development influenced the growth of a group of children over a five-year-period.

THE PROGRAM

From October, 1933, to June, 1939, some 500 families participated in a cooperative program of child development carried on in Paul Revere School, an elementary school in the city of Cleveland, by Hiram House, a social settlement. Paul Revere School had been designated as a curriculum center for experimentation with "new material and methods for the development of health, character, and citizenship" (2) Then, in the fall of 1933, under the direction of Mr. George A. Bellamy, director of Hiram House, an experimental home and community program designed to supplement the specialized program at Paul Revere School was put into effect.

Trained workers employed by Hiram House attempted to carry into the home the health, character, and citizenship program of the school. These workers attempted to teach the children of the 500 families enrolled in the cooperative program correct dietary and sleep habits. In addition, the parents of these children were shown, through home visits and through carefully organized child study groups, the importance of encouraging their children to maintain regular routines and schedules compatible with generally accepted health standards. Trained leadership in play and recreation was also provided by Hiram House, and further, all children were encouraged to accept increasingly greater responsibility for work after school hours.

Through the development of better sleep and dietary habits, through the acquisition of more satisfactory leisure-time activities, and through acting as a connecting link between the home and the school, it was hoped that better home and school adjustments would be effected, that the general physical developments of the children would improve, and that social relationships would show evidence of maturing personalities.

RELATED STUDIES

In 1935 a study by Chew (1) presented a detailed analysis of one-week records of foods eaten and the number of hours slept by 1,275 Cleveland school children. From these records, which were collected in four public schools, 8,002 dietary records and 6,337 sleep hours were studied and compared with accepted standards. Chew found that, when the Cleveland records were compared with accepted diet and sleep standards for school children, in both instances the scores of the Cleveland children participating in the study were appallingly inadequate.

* Acknowledgment: The author wishes to acknowledge her indebtedness to Mr. P. D. Graham who was largely responsible for directing the collection of the data on which this study is based, to the Cleveland Foundation whose financial assistance made the project possible, and to the Cleveland Board of Education for their splendid cooperation.

The concern of this particular study was largely with the *need* for an educational program such as that evaluated in the present study.

A second study by Conner, Hawkins, and McCarty (2) then followed, attempting to determine the effectiveness of the *first year* of the experimental program at Paul Revere School. In this later study the populations of two schools were compared on the basis of attendance records, food and sleep scores, scores on the Haggerty–Olson–Wickman Behavior Rating Schedule, pupil progress in school, and the frequency of indoor and outdoor activities and of home responsibilities.

The authors concluded that the initial phase of the work at Paul Revere School in which Hiram House had "begun to experiment tentatively to develop a program of social service . . ." (2) seemed to have been effective in providing for the "uniform development of character and school citizenship . . . " (2) The Conner, Hawkins, and McCarty study differed from the present study, however, in that the interpretations arrived at were based on data collected only from school records and from diaries kept by the children enrolled in the two schools, and further, only the first year of the experiment was studied in contrast with the five-year period surveyed in the present investigation.

Subjects

The 500 families participating in the child development program were in the fifth-tenth economic group in the city of Cleveland (3), the average income for the group being about $30 per week. Most of the fathers of the group were skilled workmen, and, during the period under survey (1933 to 1939), most of them were employed, fewer than 20% being on governmental relief.

A total of 1,149 children were involved in the program, the number of boys and girls being approximately equal. The age range was from 2 to 10 years of age, the median age at the time of enrollment being 5½ years. All of the children and practically all of the parents were American-born, the latter being of various nationality extractions.

Inasmuch as the entire child development program was designed specifically for service rather than research purposes, it has not been possible to include all of the cases in all of the comparisons given in this study. The average length of enrollment was 14.2 months, some families, of course, remaining active for much longer periods of time. In consequence, the case records of all families do not yield exactly the same quantity and kind of data.

Reliability of the Data

The items of comparison included in the present study *en bloc* represent the pooled judgments of the Hiram House counsellors, the public school psychologists, the public school teachers, the school physician, and, finally, the mothers of the families participating in the experimental program. In each instance the scores and estimates have been the result of careful and, as nearly as possible, impartial judgments. Practically all of the estimates were given with no knowledge of their probable use in evaluating the success of the experiment, since, in actuality, the problem of examining any possible changes in the subjects developed *after* the data had been collected. The estimates made by the mothers and the counsellors were originally used largely as teaching devices. The measurements obtained from the school records of the school-age children of the 500 families constituted a regular part of the school routine, and, further, these records were in existence *before* the decision had been made to evaluate the experiment at the expiration of the five-year period for which the cooperative arrangement between Hiram House and the Cleveland Board of Education was originally scheduled.

It would seem, then, that if all of the data evidenced favorable trends, the reliability need not offer obstacles in the way of interpretation.

Results

In general, the method of evaluation employed throughout the study has been to compare the scores and records at the opening of the cases, i.e., at the time of enrollment in the experiment, with the last scores and records available at the closing of the cases. All cases active in June, 1939, were arbitrarily closed.

For the sake of convenience in tabulating the data, scores at the opening of the cases have been termed the "before" ratings, and the scores at the closing of the cases have been termed the "after" ratings. Further to

facilitate the presentation of data, the scores and records have been divided into three general classifications: (1) those for which no comparative standards have been available, (2) those which might have been influenced by increased chronological age, and (3) those not likely to have been influenced by increased chronological age and for which no comparative standards or control records have been available.

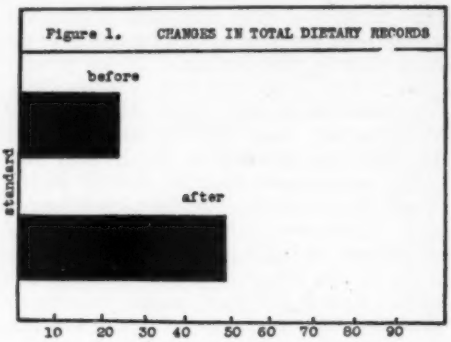
TABLE 1

DIETARY RECORDS IN TERMS OF THE NUMBER OF HELPINGS RECORDED FOR CHILDREN OF THE EXPERIMENTAL GROUP AT THE OPENING OF THE CASES AS COMPARED WITH THE NUMBER OF HELPINGS RECORDED AT THE CLOSING OF THE CASES

Dietary Items	Percentage of Cases Standard or above "Before"	"After"	Difference Between Ratings
Raw Vegetables	13.2	33.5	20.3%
Bread	41.1	58.5	17.4%
Eggs	13.3	23.3	10.0%
Butter	43.0	51.4	8.4%
Cooked Vegetables	55.2	63.0	7.8%
Meat	67.9	75.6	7.7%
Milk	32.9	49.4	6.5%
Cereals	19.3	24.7	5.4%
Fruit	47.5	52.6	5.1%
Potatoes	70.0	68.6	— 1.5%

Scores for which comparative standards have been available.—In Table 1, 370 "before" and 340 "after" dietary records are compared. As shown in the table, on nine of the ten items listed, the percentage of cases rating standard or above standard at the closing of the cases was greater than the percentage of cases rating standard or above at the time of enrollment in the experimental program, the greatest change being indicated in the consumption of raw vegetables. The smallest change, and the only negative one, was indicated in the difference between the initial and final ratings of the potato-consumption of the group. Since positive changes in other energy-producing foods seemed to have been brought about during the period, this small negative change may be regarded as inconsequential.

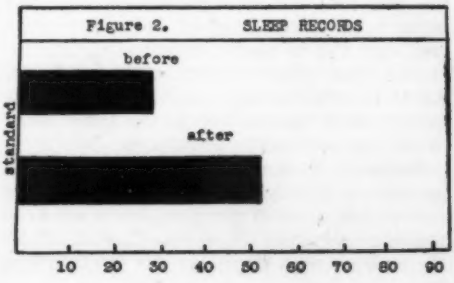
A further analysis of the dietary records of the experimental group may be seen in Figure 1. Here the comparison of "before" and "after" ratings has been in terms of the total dietary records, i.e., for this particular comparison each record was examined to determine the extent to which the total caloric content of the diet equalled or exceeded the standard. When 336 "before" records were compared with 336 "after" records, it was found that 28% more of the total dietary records were standard or above at the conclusion of the five-year-period.



The 331 sleep records represented in Figure 2 indicate a change almost equal to that shown in the total dietary records, 23% more of the cases being at or above standard on the "after" rating. The sleep records, of course, consisted of parental reports of the number of hours of sleep regularly obtained at the opening of the cases as compared with the number of hours of sleep regularly obtained at the closing of the cases.

One direct means of estimating the effects of the reported changes in dietary and sleep habits is to be found in weight changes, making allowance, of course, for the increase in height for the period under survey.

To what extent, then, did the weight of the children of the experimental group change? When 238 "before" and 224 "after" ratings by the school physician were compared, it was found that the percentage of children above the standard weight for their



46.0%, and that the percentage of children age and height changed from 38.2% to below the standard weight for their height was reduced from 47.9% to 43.8%. These weight changes would seem to offer some substantiation for the better dietary and sleep records reported for the group.

Scores corrected for age.—The development of mature personalities has been stated as one of the objectives of the experimental program at Paul Revere School. The assumption of responsibilities and getting along with other children have been regarded here as two indices of such development. If the experimental group evidenced a decrease in the percentage of children listed as having no companions or one companion, if the percentage of children reported as having three or more companions increased, and if an increase in the percentage of cases reported as having assumed a greater number of self-care responsibilities and responsibilities to the home as well, could be demonstrated, there would seem to exist some means of measuring the extent to which the children of the experimental group were developing mature personalities.

Unfortunately, standards against which these records might be checked are non-existent at the present time. Furthermore, it was not possible to obtain similar records for control cases. Since these two checks on the reported data were not possible, and since increased chronological age alone might account for any changes reported for the period of the study, it was necessary to correct the data for age.

The records were divided into age groups of 14 months each. Within each age group, then, the records were ranked according to the length of time each child was enrolled in the program. Next, each age group was divided into approximately two equal parts: (a) those who had spent more, and (b) those who had spent less time in the program than the median child for the age group. All of the a-groups were then combined into a total Group A and the b-groups into a total Group B. In this manner Group A and Group B were matched for age, and the differences found between the scores of the two larger groups might then be attributed to some factor or factors other than increased age.

In Table 2 the findings reported for the number of companions of the participating children are shown. On all three comparisons the change favors Group B, consisting of those children enrolled more than 14 months in the activities of the program. The percentage of children reported as having three or more companions increased from 51% to 75%, indicating a change of 24%. On the other hand, the percentage of children reported as having no companions or only one companion was reduced, in the first instance from 8% to 4% and in the second instance from 14% to 9%.

TABLE 2

CHANGE IN NUMBER OF COMPANIONS REPORTED FOR 377 CHILDREN PARTICIPATING IN THE EXPERIMENTAL PROGRAM AT SCHOOL X

Items Measured	Group A (Enrolled less than 14 months)	Group B (Enrolled more than 14 months)	Difference
Three or More Companions	51%	75%	24%
No Companions	8%	4%	— 4%
One Companion	14%	9%	— 5%

Responsibilities have been analyzed in terms of those duties involving self-care and those involving service of some kind to the home. Under the former division, 312 records consisting of such activities as washing, bathing, combing hair, brushing teeth, changing clothes, taking care of own possessions, and earning spending money were analyzed. The latter division, consisting of the home responsibilities reported for 310 cases, comprised such activities as running errands, washing dishes, cooking, table duties, cleaning house, washing and ironing, cleaning the walks or porch, taking care of the lawn, garden, furnace, or ashes, and caring for younger children in the family.

When the data were analyzed it was found that 67% of Group A cases reported 18 or more self-care responsibilities as compared with the 72% reported for Group B, and 18% of Group A cases reported 18 or more responsibilities to the home as compared with the 24% reported for Group B.

The indoor and outdoor play activities of the experimental group were measured in terms of the number of activities, no estimation of the amount of time spent in either activity having been made.

Indoor play activities consisting of such items as reading, listening to the radio, at-

tending motion picture shows, participating in various types of organized and unorganized play in addition to vehicular play, conversation, and observation, averaged 9 activities for both Group A and Group B.

With the exception of the first three items listed under indoor play activities, the activities reported for outdoor play activities were essentially the same as those given above. The latter data, however, also included various field sports. When the outdoor play data were analyzed in terms of A and B Groups, the average reported were the same, both indicating 8 outdoor play activities.

These findings on play activities suggest that the experimental program at Paul Revere School did not affect to a measurable degree, at least, the number of play activities of the participating children. However, since no provision for the estimation of the actual amount of time spent in either kind of activity was made, it is not possible to decide conclusively the actual effect the program might have had on the length of time the children engaged in indoor or outdoor play.

Scores not likely to have been influenced by increased chronological age and for which no comparative standards or control records have been available.—Fifteen items are represented in this section of the study. Some of the ratings such as scholarship averages, probable learning rates,[1] school attendance, and the physical ratings were obtained from the school records. The remaining items, comprised of personality ratings, estimates of physical appearance, neatness of dress, family relationships, regularity of bedtime, and evaluations of the home environment, were obtained from estimates made by members of the staff of the Hiram House staff.

In Table 3 are presented all of the negative or undesirable traits on which ratings were made at the opening and again at the closing of the cases. On these items, of course, a reduced score, i.e., a lower "after" rating, indicates a favorable change.

All of the items listed in Table 3 are self-explanatory, and all, it should be noted, indicate small but favorable changes.

Some positive or desirable items of comparison have been grouped in Table 4, and,

[1] Probable learning rate (P.L.R.) is the term applied in the Cleveland Public Schools to the IQ resulting from a group intelligence test.

TABLE 3

NEGATIVE ITEMS ON WHICH "BEFORE AND AFTER" RATINGS OF CHILD STUDY GROUP FAMILIES WERE MADE

Items Compared	No. of Cases Used	"Before" Rating	"After" Rating	Change in Rating
Emotional Maladjustment				
(a) Between Mother and Child	477	32%	27%	5%
(b) Between Both Parents and Child	477	10%	6%	4%
Physical Deficiencies				
(a) Defective Vision	155	6%	.01%	5.99%
(b) Defective Tonsils	157	20%	6%	14%
(c) Defective Teeth	154	58%	16%	52%
(d) Usually Sick	430	21%	17%	4%
Unkempt Appearance	411	16%	13%	3%
Pessimistic Home Attitude	485	20%	16%	4%

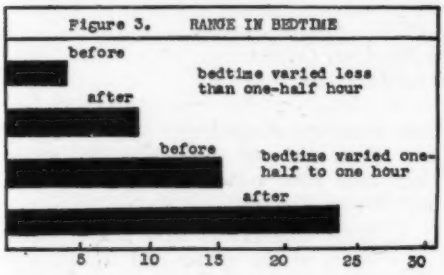
TABLE 4

SOME POSITIVE ITEMS ON WHICH "BEFORE AND AFTER" RATIONS OF CHILD STUDY GROUP FAMILIES WERE MADE

Items Compared	No. of Cases Used	"Before" Rating	"After" Rating	Change in Rating
Physical Defects Remedied				
Ear Defects Corrected	155	0%	2%	2%
Tonsil Defects Corrected	157	3%	14%	11%
Teeth Defects Corrected	154	0%	44%	44%
Healthy, Robust Appearance	432	79%	82%	3%
Neatness of Appearance	411	84%	87%	3%
Well-balanced Emotional Adjustment	477	59%	60%	10%
Cheerful Home Attitude	485	38%	41%	3%

here again, all of the items show desirable changes.

The extent to which a regular bedtime had been established for the group was estimated, and in Figure 3 the results are shown. When



ratings at the closing of the cases were compared with ratings at the opening of the cases, it was found that the percentage of children whose bed-time varied less than one-half hour at the initial rating was 4% as compared with 10% at the final rating, while the percentage varying only one-half hour to one hour increased from 16% to 24%.

Summary

The records of 500 families enrolled between October, 1933, and June, 1939, in an experimental program conducted cooperatively by Hiram House and the Cleveland Board of Education were analyzed. These records, consisting of the pooled judgments of the school psychologist, the school physician, the public school teachers, and the counsellors of the Hiram House staff, yielded two ratings used here to measure the extent to which the children involved in the study were affected by participation in the program. When the initial rating was compared with the final rating it was found that:

(1) The dietary and sleep records of the children indicated a considerable favorable change. Weight changes offered some substantiation of the improvement in dietary and sleep records.

(2) Social development as measured by the number of companions showed a definite improvement.

(3) A favorable change seemed to be indicated in character growth as measured by the assumption of increased responsibilities to both self and home.

(4) Play activities, indoor and outdoor, were not affected by the program when the number of activities alone is used as the basis of comparison. It is quite possible that this result has been affected by: (a) the fact that the time-element was not measured, and (b) the fact that the increase in responsibilities would perhaps allow little opportunity for a definite increase in the number of play activities.

(5) Physical defects showed a small but definite decline.

(6) A sizeable favorable change was indicated in emotional adjustment.

(7) Home attitudes showed small but definite favorable changes.

Bibliography

1. Chew, Helen. *Children's Food, Sleep, and Activities.* Cleveland, Ohio: The Hiram House Press, 1935.
2. Connor, W. L., Hawkins, G. C., and McCarty, K. A. "The Cleveland Experiment in the Development of Health, Character, and School Citizenship," *Journal of Educational Research,* XXXII (September, 1938), 23–34.
3. Green, Howard Whipple. *Sheet a Week,* VII (September, 1939), 1–2.

A FRAME OF REFERENCE FOR THE ORGANIZATION AND ADMINISTRATION OF CHARACTER EDUCATION IN THE ELEMENTARY SCHOOL

ISIDORE KAYFETZ
Principal, Public School No. One, Queens
New York City

FOREWORD

This attempt to formulate suggestions for the improvement of the organization and administration of character education in the elementary schools of the City of New York is an outgrowth of the analytical study of two recent surveys of character education conducted in the local educational system. The first of these (1: pp. 100–145.) is a summation of the answers made by heads of elementary and junior high schools to the following questionnaire:

"1. What phase or phases of character education are carried on within your school?
2. Outline briefly the salient features of your program.
3. What element or part of this program has been most effective?"

No attempt was made to determine the validity of the claims of the various principals and to measure the success of the different programs.

The second survey (4) was part of a broad, comprehensive study of supervision, being a summation of the answers made by the heads of 99 elementary schools to the problem:

"44. Describe briefly any organized special character training or morals and manners program planned by you and now effectively functioning in your school."

In both surveys, particularly in the one mentioned first, the committees were aware of the limitations of inquiries based upon broad general questions designed to ascertain chiefly what the existing procedures are concerning the organization and administration of character education. At best, such studies should be considered as exploratory surveys from which certain problems may emerge for more thorough and more scientific study. They may shed some light on such questions as:

1. What conceptions have school supervisors of the nature and functions of character training?
2. What are their views as to the aims and objectives of character training?
3. What curricular programs and what materials, activities, and agencies do they employ in their efforts to achieve the objectives they have in mind?
4. What principles and techniques of teaching, of learning, of organization, administration, and supervision are they using to carry out their programs of character education?
5. To what extent are the outcomes of these programs valuable and effective?
6. What problems for further investigation are revealed by these exploratory surveys?

The problems of the organization and administration of character education are the most important, as well as the most difficult and complicated in the whole field of education. They are important because the essential aim of education is the modification of personal behavior in consonance with social purposes and ideals; difficult, because there is scarcely any scientific knowledge in the field, compelling us to resort to conflicting theories in a highly controversial area; complicated, because conduct, through which character is moulded as well as manifested, is the outcome of the interplay of the energies of complex personalities reacting to the situations presented by our intricate natural and social environments.

Despite much activity in research in the field of character education in recent years, scientifically valid contributions are very

merger. They are founded, chiefly, upon questionnaires, tests, and statistical studies, rather than upon experimental methods. (10: pp. 79–179; 30). We are, therefore, obliged to draw upon moral philosophy, educational philosophy, psychology, sociology, and ethics, for most of our basic assumptions concerning the problems of character education. These problems are fundamental ones, involving the original nature of the individual, the moral nature into which we wish to mould it, the structure of society and its influence on character development, the notions of the good life that the race has evolved, as well as the strictly educational problems dealing with the aims and objectives, the agencies, the materials and activities, the organization, administration and methods of character training.

It is the considered judgment of the present writer that a critical study of the two recent surveys of character education mentioned above, in the light of the viewpoints and basic assumptions of social, moral, and educational philosophy, and the most significant conclusions of the recent literature of character education, would clarify our understanding of the fundamental problems involved, giving added meaning to the knowledge derived from a sampling, by questionnaires, of the opinions of some school supervisors as to the best methods of organizing and administering the program of character training in the New York City elementary schools. The formulation of a practical program cannot be expected to grow out of the present critical study. However, a description and an interpretation of the results of the two surveys from the standpoint of a few broad, basic assumptions and principles, should illuminate some of the fundamental issues of character education, and furnish help and guidance to supervisors and teachers in their practical endeavors in that field. Therefore, the main results gleaned from the two surveys will be considered according to the following plan:

1. Introductory considerations.
2. Presentation of the results of the two surveys.
3. Their interpretation and appraisal in the light of some broad underlying assumptions.
4. The formulation of a few general guiding principles.
5. The educational implications of these principles and their application to the problems of the organization and administration of character training in the elementary schools of New York City.

I. Introduction

A. The Need for Reconstructing Our Fundamental Positions

Despite the many profound changes in the physical, the biological, the cultural and the social foundations of human life from primitive times to the present, it has always been the chief aim of the social group to mould the character and regulate the conduct of the individual to promote the common welfare. At different stages of human development, conduct has been controlled in accordance with custom and tradition, in conformity with laws of divine or of human origin, or pursuant to intelligently directed voluntary choices. As social groups became more numerous and complicated, education became increasingly important as an agency for transmitting and transforming the social heritage to enable man to make the adjustments required for meeting the changing needs and ways of communal living. Educational development, though frequently lagging behind, to a large extent, kept pace with changes in social organizations, institutions, and culture. New conceptions gradually emerged concerning the aims and values of life, the goals of human existence, the nature of the individual, the nature of society, and the functions of education. At all times, however, despite fundamental differences in the actual rules of conduct, in their interpretation and application, and in their framework of basic ideas, one of the foremost tasks of the school has been the development of good character.

In the early history of our country the fundamental aim of character education was envisaged largely in terms of living in accordance with the doctrines of some organized religion, the whole program of education being built around that purpose. The transition of our country from a predominantly agricultural society to a predominantly commercial and industrial one, has necessitated a reexamination of every aspect of corporate life, to the end that we may build a better educational system for a better American social

order, and, let us hope, as the outcome of the present titanic struggles, a better world fellowship dominated by intelligence and goodwill rather than by force and brutality. This, naturally, involves the constant re-examination and revision of our basic conceptions of human values, of educational values, and of character education, for, it must be remembered that not all change is progress. Some changes are neither good nor beneficial. The movie, the radio, the automobile, the airplane, and numerous other material inventions have brought much good, but also much evil. Technical progress does not necessarily bring social progress. It is the use to which we apply our material and intellectual resources that counts for or against human progress. The purpose is moral only to the extent that our resources are used to promote the common welfare. Upon education, more particularly upon character education, falls the responsibility of developing individuals possessing the good-will and the intelligence to direct the use of all material, intellectual, and social inventions towards good purposes, otherwise, as the present world cataclysm amply proves, they will be used largely for purposes destructive of the highest human and spiritual values. The democratic moral and educational ideal has never been put more clearly and simply than by John Dewey: "Learn to *act* with and for others while you learn to *think* and to judge for yourself."

The pressing need of the hour is hard and serious thinking to clarify certain fundamental problems that we thought we understood, but which the present world chaos proves we fail to comprehend. What is education, what is character, what must we do to insure that education in general and character education in particular, may become realities rather than mere pious wishes? (12: pp. 115–116.)

II. The Problems of Character Education

An overview of the essential problems of character education should clarify our understanding of any plan for the improvement of its organization and administration. White and Macbeath regard the following as the main problems of character building (34: p. 10.)

1. Man's original nature or native endowment.
2. The moral nature into which we wish to transform the original nature.
3. The extent to which the original nature can be moralized.
4. The principles underlying the growth and development of character.
5. The principal agencies and organizations engaged in character training, and the part each should play in the process, especially the school.

Much of the conflict, confusion and ineffectiveness in character training is due to lack of clear understanding of such basic questions as the following:

1. What is character?
   a. How does it depend upon the natural and the social environments in which the individual lives?
   b. How does it depend upon the native endowment of the individual?
   c. What is its relation to the changing concepts of the good life that man has developed through the ages?
2. How does character grow and develop from infancy to maturity?
   a. What are the laws and principles of character development, as far as we know them?
   b. How can we apply our knowledge of these laws and principles to the prediction and control of the process of character development?
3. What agencies and organizations are concerned with character development, what part should each one play in the process, and how can their efforts be coordinated more effectively? More particularly, what contributions can be made by the family, the community, industry, the church, the press, the movies, the radio, art, science and invention, and other agencies?
4. What is the special responsibility of the school regarding character education? This involves a consideration of such matters as:
   a. The major aims of character education.
   b. The specific objectives of character education.
   c. The relation of the curriculum to character education, including,

(1) Subject matter.
(2) Co-curricular activities.
(3) Activity program, or experience curriculum.
(4) How do the trends in curricular development affect the problems of character education?

d. Methods and techniques of character education.
(1) In the classroom.
(2) Through organization, administration and supervision of the school.

e. Appraising the outcomes of character education.

It may be objected that a consideration of all the foregoing problems makes the study of character education a difficult and complicated task. Of course it does. Nothing is ever gained in the study of a complex problem by attempting to oversimplify it. By leaving out many significant elements you change the whole nature and meaning of the problem, and, consequently fail to solve it. While a consideration of the question of what teachers and supervisors do to train children in manners and morals may shed some light on character training, such a limited survey will illuminate only one small corner of the field. For a fuller, deeper understanding, we should also inquire into such essential problems as, what is character; why do we follow certain procedures in character training; what are the outcomes of our efforts; what are the values of these outcomes, and how shall the values be appraised?

Much could be learned by a thorough analysis of the two surveys of character education in relation to all the major problems involved. This is a larger task than the scope of this paper will warrant. However, the writer will indicate the broad outlines of such an analysis in the light of several significant underlying assumptions. These are put forward not as a fixed, rigid platform that a "program" of character education must stand upon, but rather as an elastic springboard for a dive into the turbulent waters of character training, in the hope that clear heads, strong limbs, and stout hearts may help us reach our destination.

1. The essence of character is consistency of conduct in accordance with an integrated system of habits and attitudes, within a framework of regulative principles and ideals, hence a plan of character education must grow out of a coherent system of values. This is to say that character education must be grounded upon a clear, consistent and comprehensive social and educational philosophy.
2. In our American social order character education must be envisioned from the point of view of democratic principles and practices, both as the avenue of approach and as the end. Our major concern must be the training of our children for the democratic way of life, in the broadest and most inclusive sense, embracing every sphere of human activity, and every relationship within the school. Such a broad, comprehensive purpose, encompasses in a unified whole, the aim, the method, and the goal of character education.
3. The essence of moral responsibility is freedom of choice and action in accordance with ideals of conduct which the individual has approved, accepted and made his own.
4. There is no essential conflict between the pursuit of individual happiness and welfare and the promotion of the happiness and welfare of the group. The two are complementary, the individual fulfilling and realizing himself in working for the common good.
5. Democracy and moral responsibility are not separate subjects referring to independent bodies of experience to be superimposed upon or added to an overcrowded curriculum and to be learned in isolation from other phases of school experience, mainly in verbalized forms. They are as broad and all-inclusive as life itself. Their spirit and substance should permeate and pervade every aspect and every activity of school life and they should be learned as much as possible through action and practice.
6. From both the individual and the social points of view, the essence of moral responsibility is "integration of values—doing the best thing in each situation." (10: pp. 55–59.) For the individual in relation to himself it

means doing the best thing, in all situations in which he finds himself in the pursuit of the aims of self-realization. From the social point of view, it means doing the best thing in all situations in which the individual finds himself as a member of various organized groups, in performing the duties and discharging the responsibilities involved in all his social relationships.

Having now laid the groundwork, let us consider the two recent surveys of character education.

III. Assistant Superintendents' Survey of Character Education

A. *Aim:* The aim of this survey was to find out what the New York City schools are doing in character education. (1: p. 101.)

B. *Method:* The method of obtaining data was the submission to principals of the following questionnaire:

"1. What phase or phases of character education are carried on within your school?
2. Outline briefly the salient features of your program.
3. What element or part of this program has been most effective?"

C. *Results:*

1. *Mode of presentation:* The answers were studied and presented only in descriptive terms, as they evidently did not lend themselves to objective, quantitative analysis. Thus, the summary presents a composite picture of all parts or phases of the character education programs in use in all schools, the most commonly used procedures being emphasized.
2. *Conception of the Nature of Character:* The majority of the principals believed that the predominating aim of school organization and practice should be "the development of our boys and girls into wholesome and well-integrated personalities." (1: p. 102.) Desirable understandings, attitudes and ideals must be developed which will find expression in tendencies to right conduct in the situations of daily life, with the school assuming the task of developing and strengthening such understandings, habits, attitudes, ideals and dispositions through providing ample opportunity for their practice in the social life of the school community.

   Various traits of character are emphasized including social, individual, civic, and specific "school" traits peculiar to the conditions of particular schools.

   There is very little recognition in the survey of the dynamic nature of character, and of the social environment and of how they affect each other. This points to the need for a more thorough study of the psychological and the sociological bases of character education. (30: No. 2.)
3. *Cooperation with other agencies:* There is very little recognition in the report of the part played by agencies other than the school in moulding character, such as those mentioned on page 4 of this article. Brief mention is made of cooperation with parents, with Parents' Associations, with religious organizations, with health and philanthropic organizations. There is need for closer cooperation between the school and other agencies. (16)
4. *The School and character education:*
   a. *Underlying philosophy:* No attempt has been made to state a point of view, to formulate the underlying assumptions of a social and an educational philosophy, or to bring character education into relationship with such underlying assumptions. This shortcoming is recognized in the survey when it points out that the results reveal varied and conflicting theories, practices and procedures, which may be justifiable and desirable insofar as these reflect differences among pupils, types of schools and community problems, but there is no justification for lack of agreement on fundamental points of view and philosophy. With this the present writer is in hearty agreement, having anticipated the recommendation in the survey that there is "a need for determination of the basic as-

sumptions underlying a broad outline of character education", by suggesting a number of "basic assumptions" on pages 5 and 6 of this article.

b. *Character education and the curriculum:*

(1) The curriculum is regarded as the determining factor in character education, particularly if it is conceived in the broad sense of all the experiences the pupils engage in under the influence of the school, recognizes the needs, purposes, and interests of each individual child; selects materials and activities suited to each child; provides for remedial work; uses materials and activities related to life.

(2) *Subject Matter:* The character-building implications of various types of subject matter are stressed, particularly the social sciences, the physical sciences, the biological sciences, health education, homemaking, English and the appreciation subjects.

(3) *Extra-Curricular Activities:* Such extra-curricular activities as club activities, assembly programs, character training organizations, recreational activities, sports, athletics, and others are regarded as valuable for character training.

c. *Methods and Techniques:*

(1) *In the classroom:* The following methods were regarded as valuable in character training:

(a) Those which stimulate the imagination, arouse desirable emotions and moral expression.

(b) Socialized recitations.

(c) "Activity program."

(d) Creative expression.

(e) Cooperative learning.

(f) Direct instruction by means of moral lessons, mottoes and slogans, readings in character readers, etc.

(g) Indirect or incidental instruction.

(2) *Through the organization, administration, and supervision of the school:* Various features of organization, administration, and supervision contributed to character education, notably:

(a) School situations and routine.

(b) Development of social consciousness.

(c) Building school spirit.

(d) Creating a democratic spirit.

(e) Emphasizing high ideals.

(f) Teachers and supervisors discussing and agreeing upon the meaning of good character and personality integration.

(g) Teachers and supervisors setting good examples.

(h) Teachers with fine personalities.

(i) Classification of pupils in special classes according to their individual needs.

(j) Pupil participation in school management.

(k) Individual guidance.

d. *Appraising the Outcomes of Character Education:* The survey does not attempt to determine the validity of the procedures reported, nor does it discuss any standards for appraising the effectiveness of the character training program of any of the schools.

e. *Problems for Further Study:* The survey raises a number of questions, some of which it recommends for further study. The questions are as follows:

(1) Does the daily life of the pupils in their schools help to make citizens possessing the attitudes and ideals of cooperation required for the building of a better world order based upon mutual understanding and interdependence?
(2) Which program of character education is more effective, the direct or the indirect?
(3) How can we bring about a wider recognition and a more effective use in all schools of daily routine, the curriculum, the methods of instruction and extra-curricular activities as potent means of character training?
(4) What is the effect of homogeneous grouping and heterogeneous grouping on character education?
(5) Do modern philosophy and modern psychology justify the extensive use of school incentives?
(6) What is the effectiveness of character training organizations on character education?
(7) Which extra-curricular activities are of greatest value for character education?
(8) What interest should the school take in the pupil's learning outside the school and how shall such learning be coordinated with the work of the school?
(9) How shall a better mutual understanding be attained between social agencies and the school in dealing with the problems of assisting the children whose behavior is anti-social?
(10) How can the attitudes that characterize mental hygiene be secured in all the schools?
(11) How can the schools insure the development of good teacher personality?

Finally, the survey recommends a "more thorough and a more scientific study" of the problems of character education.

IV. Survey of Supervision

A. *Aim:* The purpose of question number 44 of the survey of supervision was to ascertain what "organized special character training or morals and manners program" was planned and was effectively functioning in the 99 elementary schools involved in the investigation. (4)

B. *Method:* The method for obtaining data was the submission to the heads of 99 elementary schools of the following questionnaire.

"Describe briefly any organized special character training or morals and manners program planned by you and now effectively functioning in your school."

C. *Results:*

1. *Reports:* Ten principals or approximately 10% of the total number reported that they had no organized programs. These replies may be explained by the apparent misunderstanding on the part of these supervisors of the meaning of the question. They may have believed that what was requested was a description of a plan of direct character instruction in the form of a course of study or series of lessons.

The remaining 89 supervisors made reports of varying length and quality, which were analyzed, digested, classified and recorded in the form of an extensive table, setting forth the activities and procedures employed under the following headings:

I. Direct Measures
   A. Developmental
   B. Corrective

II. Indirect Measures
   A. Developmental
   B. Preventive

In connection with each procedure, the schools using it were designated by symbol, total number indicated in percentages.

This table has been separated into two smaller ones, Table I indicating the direct measures and Table II the indirect measures. The symbols indicating the schools using the various measures have been omitted.

TABLE I

DIRECT MEASURES USED BY VARIOUS SCHOOLS IN ORDER OF FREQUENCY

	Schools Using Them	% of Total Number
A. Predominantly Developmental Measures		
1. Direct character training through series of planned lessons	------	13%
2. Training in habits of courtesy by example and precept	------	12%
3. Development of proper habits and attitudes	------	12%
4. Assembly talks by principals and teachers	------	12%
5. Mottoes, slogans and proverbs	------	9%
6. Class discussion of behavior problems	------	8%
7. Posters and class records for good citizenship	------	7%
8. Lectures by outside speakers	------	6%
9. Moral lessons arising out of actual school and community behavior problems and situations	------	6%
10. Readings on manners and morals	------	6%
11. Principles of mental hygiene	------	5%
12. Setting up standards of socially approved conduct by teachers and pupils cooperating	------	4%
13. Special circulars on manners and morals	------	3%
14. Etiquette club	------	3%
15. Assembly forums on standards of behavior	------	3%
16. Direct teaching of manners	------	3%
17. Religious training through urging pupils to attend religious institutions	------	3%
18. Stressing ideal of cooperation	------	2%
19. Humane education	------	1%
20. Hutchins' childrens' morality code	------	1%
21. Transfer of training through evolving and applying generalized ideas of action	------	------
a. Providing rational basis for action	(12%)	3%
b. Honesty and fair play	------	2%
c. Self-control	------	2%
d. Social participation	------	1%
e. Pupil initiative	------	1%
f. Responsibility to group	------	1%
g. Connecting dissatisfaction with failure	------	1%
h. Connecting satisfaction with success	------	1%
22. Pupils' committees to evaluate behavior	------	1%
B. Predominantly Corrective Measures		
1. Individual corrective measures	------	6%
2. Self-rating of character traits by pupils	------	4%
3. Check list of character traits		
a. For all pupils	------	1%
b. For problem pupils	------	1%

An analytical study of Tables I and II reveals the following general facts:

1. 10% of the schools reported that they had no programs, due perhaps to a misunderstanding of the question.
2. Many of the schools used a combination of direct and indirect procedures with the indirect preponderating.
3. The highest average percentage of procedures were the indirect, preventive ones, such as: school monitors, safety patrol, sanitation squads, etc. This seems to indicate that the schools regard, or find it necessary to use, measures to prevent undesirable behavior, as the chief elements in their programs of character training.
4. Next in order, in terms of average percentage, but largest in number and variety of procedures were the indirect, developmental ones, such as: assembly programs, school incentives, subject matter, pupil participation in school government, character organizations, cooperative activities of pupils, functional teaching of manners, thrift accounts, cooperation between school and social agencies, etc. This seems to indicate that the indirect or incidental method of character training is considered more effective and is used more than the direct method.
5. Direct, developmental procedures rank third in average percentage, but second in the number and variety of procedures

TABLE II

INDIRECT MEASURES USED BY VARIOUS SCHOOLS IN ORDER OF FREQUENCY

II. Indirect Measures

A. Developmental

1. Assembly programs	30%
2. Competitive awards (banners, prizes, etc.)	28%
3. Moral implications of various types of subject matter	24%
4. Pupil participation in school management	18%
5. Club programs	15%
6. Character organizations	12%
7. Cooperative activities of pupils	10%
8. Functional teaching of manners	9%
9. Thrift (Savings accounts)	9%
10. Cooperation between school and social agencies	8%
11. Social welfare activities	7%
12. Cooperation with Parents' Associations	7%
13. Extracurricular activities	7%
14. School honor societies	5%
15. Attendance contests	4%
16. Cooperation between school and home	4%
17. Teacher training through conference and other means	4%
18. Cooperation with recreational agencies	4%
19. Play, games and athletics	4%
20. Honor Rolls	4%
21. Assembly programs conducted by pupils	4%
22. Ethical appraisal of current events	4%
23. Cooperation of school and community	3%
24. Socialized recitations	3%
25. School publications	3%
26. Cooperation between school and religious organizations	3%
27. Case method	3%
28. Excursions	2%
29. Motion pictures and radio	2%
30. Committee on character education	2%
31. Adaptation of course of study to individual pupils	1%
32. Class organized as manners and morals club	1%
33. Handling of money by children to encourage honesty	1%
34. Professional reading by teachers and supervisors	1%
35. Questionnaires to teachers as basis for setting up program	1%
36. Recommending special Radio programs	1%
37. Special honor roll for students mentally incapable of meeting regular standards	1%
38. Miscellaneous	
a. Flag Day Pageant	1%
b. Bugle Corps	1%

B. Preventive

1. School Monitors	22%
2. Safety Patrols	14%
3. Sanitation Squads	10%
4. Service Leagues	5%
5. Rules and regulations reduced to a minimum	2%
6. Avoidance of situations which might result in undesirable behavior	1%

used. When coupled with the fact that indirect developmental measures ranked highest in number and variety of procedures, it appears that the schools concern themselves largely with the *development*, by both direct and indirect methods, of what they regard as desirable understandings, habits, attitudes, ideals and dispositions making for good character. However, there is very little agreement as to what these are.

6. Lowest in average percentage, and in number and variety of procedures are the direct, corrective ones such as: individual corrective measures, self-rating of character traits by pupils, etc.
7. The different schools vary greatly in both the number and the kind of procedures they use, indicating differences in both the amount and the quality of character training in the various schools, and in its organization and administration.

TABLE III

TEN MOST FREQUENTLY USED MEASURES

Type of Measure	% of Schools
Assembly programs	30%
Competitive awards (banners, prizes, etc.)	28%
Moral implications of subject matter	24%
School monitors	22%
Pupil participation in school government	18%
Club programs	15%
Safety patrol	14%
Series of planned lessons on morals and manners	13%
Training in habits of courtesy by example and precept	12%
Development of proper habits and attitudes	12%

Table III indicates, in the order of frequency, the ten most popular procedures. A number of striking facts appear from a study of this table. The first, second and third, which are assembly programs, school incentives, and subject matter; the fifth and sixth, pupil participation in school government and club programs, are all indirect developmental procedures; the fourth, school monitors, and the seventh, safety patrols, are indirect preventive procedures; the eighth, ninth and tenth, series of planned lessons on morals and manners, training in habits of courtesy by example and precept, development of proper habits and attitudes, are direct, developmental procedures. From this, it would appear, that for the most part, the character education program in many schools is a mixed program of direct and indirect, developmental and preventive procedures, with indirect, developmental procedures predominating and indirect preventive measures of a policing nature playing an important part. The greatest importance seems to be attached to the efficacy for moral training of assembly programs, school incentives, subject matter and school monitors.

Only a relatively small percentage of the total number of schools agree upon these most popular measures.

A study of Tables I, II, and III reveals the striking fact that 69 different specific procedures are reported, 25 of which are direct and 44 indirect. Outside of the ten most popular measures listed in Table III, there are no less than 59 specific measures each of which is used by 10% or a smaller proportion of the schools involved. This seems to indicate, even if we make allowance for differences in types of schools, in types of pupils, and in neighborhood conditions, a lack of system, coherence and unity in the character training of the elementary schools. This may be due largely, as was pointed out in the Assistant Superintendents' Survey, to the absence of a fundamental viewpoint and of an underlying philosophy of both general and character education. In this connection it may be well to consider the injunction of White and Macbeath, who emphasize that in moral education, ethical theory is essential, because, "no one, except fortuitously, can expect to mould or help produce a good moral character until he has a grasp of the ideal he wants to realize. The difficulties of realizing the ideal are great enough, even if he knows what to aim at; without a clear and definite grasp of the end, they are insuperable." (34: p. 7.)

V. CRITICAL APPRAISAL OF THE ASSISTANT SUPERINTENDENTS' REPORT AND THE SURVEY OF SUPERVISION

I. *Aims:* There was essential similarity in the aims of the two surveys. The Assistant Superintendents' survey was undertaken to find out what the New York City schools are doing in character education. (1: p. 101.) The Survey of Supervision was undertaken to "make a study of the duties performed by supervisors, the time spent in such duties, and the duties, if any, which are not being performed by them for lack of time." (4) The programs of character education being used in the 99 elementary schools constituted one of the collateral issues of the broad, comprehensive survey of supervision. While the Assistant Superintendents' survey included both the elementary and the junior high school, the report nowhere mentions what schools or how many schools were included in the study. The other survey was confined to the elementary schools and included records by means of which each of the 99 schools involved could be identified. Both surveys were concerned with studying existing programs and practices. The main questions regarding which information was sought were, "What is being done in character education and how is it being done?" Neither survey deals explicitly with such important questions as what is the essential nature of character, what are the major aims of general education, what are the major aims of character education and their relationship to the aims of general edu-

cation, what is the philosophical frame of reference for both sets of aims, why are certain procedures being used in character training, what are the values of these procedures, and how are these values determined? The Assistant Superintendents' report states that it is describing "procedures helpful in character education viewed as personality integration," (1: p. 100.) and also defines character in the same terms. (1: p. 102.) It also indicates the "need for determination of the basic assumptions underlying a broad outline of character education." (1: p. 142.)

The problem is stated more clearly in the Assistant Superintendents' survey, as is probably indicated by the fact that in the Survey of Supervision, 10% of the school heads stated that they had no programs.

It is the opinion of the present writer that the fundamental requirements in any attempt to organize and administer a "program" of character education in the elementary schools are:

1. The formulation of the basic philosophical assumptions underlying such a "program."
2. Clear conceptions of:
   a. The aims of, and objectives of democratic education.
   b. The aims and objectives of character education in a democracy.
   c. The psychological and the sociological bases of character, viewed dynamically.
   d. The functional relationship between the aims and objectives of general education and of character education.

In this paper, considerations of space, make it impossible to deal with the psychological and the sociological bases of character. However, there is a large amount of material on these phases of the problem which should be studied. (30: No. 2.) The writer will confine himself chiefly to an overview of the other fundamental problems enumerated above.

II. *Methods:* In both surveys the method used was the questionnaire, long considered a poor methodological instrument, yielding results which are neither valid, reliable, nor objective. At best, it can be used mainly for exploratory purposes, to indicate what the difficulties and problems are requiring further study by more scientific methods. The questionnaires used in both surveys were broad and general, allowing a great variety of answers. The present writer studied, analyzed, and tabulated the answers to the questionnaire on character education in the Survey of Supervision. The answers varied from the one word "None", given in 10% of the cases, to statements four pages in length, with considerable variation in both the quality as well as the length of the answers.

III. *Results:* In the Assistant Superintendents' study the results are presented in general terms in descriptive form. Practices and procedures used in character education are described without any attempt being made to indicate how widespread or how effective any given procedure is. The words "most principals", "many principals", "many schools", "some schools", "other principals", "the schools", "a few principals", "a few schools", and similar vague, general terms are used to indicate to what extent a certain procedure is being used. The total effect of the report is a composite picture of the practices used in the schools without any indication as to what phases or portions of this composite program is being used in any one school. A superficial reader may get the impression that many things are being done in character education in the New York City schools, but one cannot tell from the report how complete, adequate, appropriate, or good the character education program is in any one school. This weakness is partially recognized by the statement that "no investigation of the validity of the claims stated by the various principals and no attempt to measure the success of the different programs have been made." (1: p. 101.)

In the Survey of Supervision an elaborate table has been prepared, but not published in the report, classifying the various specific procedures used under the main headings, "Direct" and "Indirect" and the subheadings, "Developmental", "Preventive" and "Corrective." These may be defined as follows:

1. Direct—the explicit teaching of rules of conduct, good habits, desirable attitudes, ideals and dispositions.
   a. Developmental—promoting the formation of good habits, attitudes and dispositions.

b. Preventive — guarding against the formation of bad habits, attitudes and dispositions.
c. Corrective—correcting bad habits, attitudes and dispositions already formed.

2. Indirect—the implicit or incidental training in desirable habits, attitudes, ideals and dispositions growing out of the organization, administration, and supervision of the school, and the operation of the entire teaching and learning program.

The table also shows which schools and what percentage of the total number of schools used each procedure, but does not indicate how valuable and effective each procedure was. Nor does this survey show the underlying philosophies, basic assumptions, or conceptions of the nature of character, involved in the various procedures reported. One is obliged to conclude that there is little unity, coherence, or organization in the "program" of character education, which, in turn, indicates the lack of an underlying philosophy. This undoubtedly points to the need for an analysis of the problems of character education in the light of a democratic social and educational philosophy from which may emerge clear and definite aims, as a basis for the organization and administration of character education to meet the needs and purposes of American children growing and developing in the American social environment.

VI. The Organization and Administration of Character Training

A. *The Basis for Improving Character Training:* The view is here presented, that the task of improving character training in the elementary schools of New York City does not consist of the setting up of a special plan or program of character education, as if it were a separate subject, an organized body of experiences with accompanying materials and activities to be superimposed upon or added to the already overcrowded elementary school curriculum, to be learned in isolation, separate and apart from other phases of school experience. The present writer agrees with the Department of Superintendence that character education is as broad as life itself, and not a separate phase of life. Hence, if we succeed in building a good elementary school curriculum in New York City, we will have a good character education "program," for the chief aim of moral training is the enrichment of life in school and outside of school.

It must be apparent that the best social and educational theories and the finest curriculum are not self-operating. They can be effective only to the extent that the schools are organized, administered and supervised in such ways as to facilitate the greatest possible progress towards the attainment of the goals of general and of character education.

B. *The Nature and Functions of Education and The Implications for Character Education:* In a series of recent reports the Educational Policies Commission of the National Education Association has made valuable contributions to the analysis of the functions, purposes and structure of American democratic education in the light of present social conditions and problems. (18, 19, 20) The report on the purposes of education opens with a statement that every formulation of educational purposes is essentially a statement of preferences, choices and values depending "upon the judgment of some person or group as to what is good and what is bad, what is true and what is false, what is ugly and what is beautiful, what is valuable and what is worthless, in the conduct of human affairs." (18: p. 1.) The report says further, "Educational purposes, then, are a form of social policy, a program of social action based on some accepted scale of values." (18: p. 4.) Hence, every phase of education must be formulated, planned and managed in terms of American purposes and ideals. In essence, the fundamental aim of democratic education is, to provide for each individual the maximum opportunity for full and free self-realization in harmony with the maintenance and promotion of the general welfare. An examination of various recent statements of the nature and functions of education, particularly those of John Dewey and his followers, one of whom, Kefauver, has recently formulated the "fourteen points" of a contemporary educational program, has led the present writer to define education in the following terms: (25: pp. 2–8.)

> Formal education is the conscious and systematic direction and guidance of the growth and development of the desirable physical, intellectual, emotional, aesthetic

and moral capacities and potentialities of the individual, by the continuous and progressive enlargement of his experience, through functioning materials and activities, for the purpose of promoting the self-realization of a wholesome personality in consonance with the aims and ideals of social order and of social progress.

C. *Aims of Education:* As a further step in clearing the ground and building the foundation for the organization and administration of character education, we must keep in mind the major aims of general education as directives for our efforts in attaining the goals of individual and of social existence. Before doing this, however, it may be helpful to clear up some confusion arising out of the careless use of educational terminology. The terms "aims," "principles", "objectives", and "outcomes" are used synonymously by many writers, when as a matter of fact they have different meanings. The famous "Cardinal Principles of Secondary Education" (8) which have had a tremendous influence on American education since their formulation in 1918, are really not "principles", but "aims", with the exception of "command of fundamental processes", which deals with the tools of knowledge useful and necessary for the attainment of the objectives of education.

According to John Dewey, (11: pp. 117–121.) to have an aim is to act with meaning, to mean to do something and to perceive the meaning of things in the light of that intent, or in other words, acting with an aim as a foreseen end which gives direction to the activity, means acting intelligently. The goal or objective is the end toward which the activity is directed, while the outcome is the result in terms of amount or degree of progress made towards reaching the goal. (33: pp. 128–129.)

With these distinctions in mind, let us consider a formulation of the aims of education, based, with some modifications and elaborations, upon the report of the Educational Policies Commission of the National Education Association, already referred to. These are:

1. Self Realization: including physical and mental development from the point of view of mental hygiene; intellectual development, aesthetic development, moral development, integrative development, or development of wholesome personality.
2. Social Relationships: including family relationships, economic relationships, civic relationships, and other social relationships, such as friendship and general sociability.
3. Economic Efficiency and Well Being: including producer, distributor and consumer of goods and services, as well as promoter, financier and entrepreneur in economic enterprises.
4. Civic Responsibility: to local community, city, state, nation and their laws and institutions, as well as to humanity at large.

As standards of value for appraising the worth of conduct in the areas of living involved in the foregoing aims, the present writer proposes the following additional aims:

5. Moral Responsibility: to self and to others, which should pervade all life both in and outside the school.
6. Religious Responsibility: to self and to others, not conceived of in the narrow sectarian sense, but in the broad sense of religion in action, which should also pervade all life both in and outside the school.

Since the development of good moral character is so frequently stated as the most important by-product of education, whether by means of regular classroom or extra-classroom activities, it may be well at this point to consider a standard of character education set up by the Department of Superintendence in their Tenth Yearbook. (10: pp. 55–59.) From both the individual and the social points of view, the general aim of character education is "integration of values—doing the best thing in each situation." For the individual in relation to himself this standard means doing the best thing in all situations in which he finds himself in the pursuit of the aims of self-realization. From the social point of view, it involves doing the best thing in all situations in which the individual finds himself in performing the duties pertaining to social relationships, economic efficiency and civic responsibility. In both the individual and the social spheres of human conduct, the standard of values formulated above means doing

the best thing in all life situations involving moral and religious responsibility. Applying the same criterion to extra-classroom activities, the requirement is that they be pursued in such a manner as to yield the best results in personal and group satisfactions and well-being. The same criterion applies to the activities of teachers, administrators, supervisors and parents in relation to life both inside and outside the school.

The foregoing criterion of character was formulated by the Department of Superintendence in their Tenth Yearbook as the most acceptable one after a critical analysis of sixteen other conceptions contained in several hundred courses of study, books and articles. The following is a brief summary of each of the seventeen central aims of character which were critically appraised, the first sixteen being rejected and the seventeenth being accepted as the best.

1. Character as general goodness.
2. Conformity to conventional mores—doing what society expects.
3. Life in accord with the dogma of some religion.
4. Composite of many specific conduct habits capable of determination by scientific analyses of life.
5. Service of the state.
6. Social usefulness, personal self-sacrifice for the larger good.
7. Unselfish motives, love of fellowmen, desire to serve.
8. Harmonious adjustment of the personality.
9. Composite of desirable traits, virtues and ideals.
10. Self-control; inhibition of impulses in accord with rational principles.
11. Self-expression; responsibility for getting as much as possible out of one's own life.
12. Emotional maturity, objectivity, disinterestedness, intelligent living, foresight, understanding and discrimination of consequences, fair-mindedness, scientific spirit.
13. Ways of living that are aesthetically preferable, beauty.
14. Sincere action, in accord with conscience.
15. Imitation of some ideal persons.
16. Creative experience; continuous reconstruction of life.
17. The integration of values; doing the "best" thing in each situation. (10: pp. 31–59.)

D. *Objectives of Education:* The next step in the organization and administration of character education is to translate the general, directive aims of education into more definite, concrete and specific terms which will furnish guidance to the efforts of all those engaged in the teaching-learning process. This requires the analysis of our broad aims and breaking them down into more specific, subsidiary aims. The following analysis is merely suggestive of the method that is proposed. Some other analysis may be as good or superior.

1. Self-Realization
   a. Physical and mental development from the point of view of hygiene.
   b. Intellectual development.
   c. Aesthetic development.
   d. Moral development.
   e. Integrative development of a wholesome personality.
2. Social Relationships
   a. Family relationships.
      (1) Husband and wife.
      (2) Parent and child.
      (3) Brotherhood.
      (4) Sisterhood.
      (5) Other blood relationships.
   b. Economic Relationships.
      (1) Employer and employee.
      (2) Employer and employer.
      (3) Employee and employee.
      (4) Employer to public.
      (5) Employee to public.
   c. Civic Relationships.
      (1) Citizen to citizen.
      (2) Citizen to local community and its institutions.
      (3) Citizen to party organization.
      (4) Citizen to municipality and its institutions.
      (5) Citizen to state and its institutions.
      (6) Citizen to nation and its institutions.
      (7) Leadership.
      (8) Followership.

d. Other Social Relationships.
   (1) Acquaintanceship.
   (2) Friendship.
   (3) General sociability.

3. Economic Efficiency and Well Being
   a. Producer of goods and services.
   b. Purveyor or distributor of goods and services.
   c. Promoter, financier, entrepreneur in economic enterprises.
   d. Consumer of goods and services.

4. Civic Responsibility
   a. To local community and its institutions.
   b. To city and its institutions.
   c. To state and its institutions.
   d. To nation and its institutions.
   e. To humanity at large (world citizenship).

There are other educational aims which are both individual and social in their nature. These are:

5. Moral Responsibility
   a. To self.
   b. To others.

6. Religious Responsibility.
   a. To self.
   b. To others.
      (1) Your own religious group.
      (2) Other religious groups.
      (3) Persons of no religious affiliation.
      (4) Humanity at large.

The next step is to continue the analysis further by breaking down the subsidiary aims of general education into more definite and specific objectives. Since our major task is to promote the growth and development of the individual along the lines and in the directions indicated by our general and subsidiary aims, we must now get down to the individual child and formulate objectives in terms of those desirable and attainable understandings, abilities, habits, appreciations, attitudes and dispositions which, when adequately developed and properly integrated will produce persons whose conduct in each life situation, in every sphere of human activity, will be the best conduct in that situation.

To follow out the analysis in this manner it would be necessary to determine what understandings, abilities, habits, appreciations, attitudes and dispositions are required in each area of living represented by the major and subsidiary aims of education that have been enumerated above. We must constantly keep in mind, however, that a developing individual in a dynamic society, requires an active, flexible personality. This requirement is well stated in a recent report of the National Education Association, as follows:

> "Participation in our cultural resources should beget personalities who are active, not passive and inert; who are motivated by intelligently chosen purposes, not by unguided impulse from within or casual pressure from without; who are not set and rigid, but who readapt flexibly to social change and to the consequences of their own prior conduct; who express their individual differences, but who do it in ways that are cooperative and socially contributory, not self-centered and egoistic." (29: p. 12.)

The report goes on further to say that goals must be consciously pursued that will foster personal initiative, discriminating judgment and choice, flexibility of thought and conduct, individual differences, and cooperativeness.

E. *Recent Trends in Curricular Development:* Having analyzed the various major and subordinate aims of education so as to yield objectives in terms of the required understandings, abilities, habits, attitudes, appreciations, and dispositions, the next task is that of indicating the broad outlines of a curriculum that will promote the effective organization and administration of character education. Here we encounter various concepts of the nature of the curriculum, falling mainly into three general groups:

1. A group of subjects or fields of study arranged in a particular sequence.
2. The subject matter or content that is to be employed in instruction.
3. All the experiences children have under the guidance of teachers. (9: pp. 65–69.)

The third conception is gaining increasing acceptance among workers in the field of curriculum building and is the one followed in this paper. It is evident that such a broad conception of the nature of the curriculum renders the term extra-curricular activities obsolete. Fretwell recognized this when he wrote in the 25th Yearbook of the National

Society for the Study of Education, "In the more progressive schools it seems that the line of demarcation between what is curricular and what is extra-curricular is less sharply drawn than in what may be called old-line schools." (28: p. 77.) We find Jordan using the term "extra-classroom activities," (24) and Borgeson, "all-school activities" (5) and "group interest activities," (6) while McKown omits all qualifying adjectives, referring merely to activities. (26)

However, we must realize that curriculum development is in a state of transition and can be presented graphically in the manner indicated by Dr. Ferriss of Cornell University:

1	2	3	4
Pure subject matter	Subject Matter + Activities	Activities + Subject Matter	Pure Activities

First we have the extreme, traditional subject-matter curriculum in which learning is conceived of as the mastery of verbal material set-out-to-be-learned. Such programs still exist in a few schools, where either there are no other activities, or if there are they are still conceived of as extra-curricular.

In the second type, besides reading, writing, figuring and memorizing, a few additional activities may have been added in the classroom so as to make the work more interesting. These may include handwork, sewing, cooking, play and some creative activities. But the subject-matter emphasis still predominates. This is probably the condition that exists in most schools today.

In the third type, which represents the conditions prevailing in most "progressive" private and experimental schools and an increasing number of public schools, attempts are being made to work out activity programs, with subject-matter being utilized wherever and whenever necessary in the prosecution of activities. In this type of school the line of demarcation between curricular and extra-curricular is so faint as to be scarcely visible.

The fourth type is the school representing the extreme in the development of the activity or experiential program such as is being experimented with in California, particularly in Santa Barbara as well as in Virginia and in Mississippi. (22: Chap. IV.)

The movement from the subject curriculum to the experience curriculum has been described and interpreted by Hopkins as a movement towards integration passing through the following stages:

1. Subject curriculum.
2. Correlated curriculum.
3. Broad-fields curriculum.
4. Core curriculum.
5. Experience curriculum. (23: X–XIII.)

No doubt, in each of these types of curriculum, extra-classroom activities play a different part, but limitations of space prevent a detailed examination of the problem.

The problem of curriculum building is largely a matter of the selection, organization and sequential arrangement of appropriate learning situations in the various areas of living or experiencing. The importance of the situation is receiving increasing attention in recent pedagogical literature. A few examples will suffice:

> "The center of attention is to be the situation . . . Life is one situation after another, and each situation has possibilities of richer or poorer living, of greater or less integration of values." (10: p. 57.)
>
> "Different situations succeed one another . . . As an individual passes from one situation to another, his world, his environment, expands or contracts . . . What he has learned in the way of knowledge and skill in one situation becomes an instrument of understanding and dealing effectively with the situations which follow. The process goes on as long as life and learning continue." (12: p. 42.)

The term "unit" or "unit of learning" is more frequently used than the term situation as a reaction against the fragmentary nature of traditional learning of subjects as bodies of facts, principles, and processes entirely outside the individual and not springing from his interests, purposes and needs. Harap defines a learning unit as "a complete and co-

herent learning experience having a purpose which is meaningful to the pupil, accepted as his own, and which is closely related to a life situation." (22: p. 77.) instead of the terms "unit", "unit of work," "unit of learning," the present writer prefers the terms used by Melvin, "units of conduct," or "units of life." (27: p. 101.)

In accordance with the above views, an experiential curriculum is a sequence of "units of conduct" or "units of life", which are larger or smaller units of experience finding their source of unity in the learner.

Naturally, for the elementary school, the curriculum must be planned in terms of the experiences of the pupils and their individual and social purposes and needs. The simple basis of Harap's plan of "areas of living" is highly suggestive, in which life activities are grouped as follows:

1. Living in the home.
2. Leisure.
3. Citizenship.
4. Organized group life.
5. Consumption.
6. Production.
7. Communication.
8. Transportation.

If this classification is adopted as the basis of the curriculum, each "area of living" should be analyzed in the light of the understandings, abilities, habits, appreciations, attitudes and dispositions necessary for desirable conduct in the various situations arising in these areas.

That the trend of development in New York City is along the lines indicated above is shown by the recommendation contained in the recent report on the six year "activity program" experiment (3: p. 240.) which says, "In the program of curriculum revision contemplated for elementary schools, incorporate such content and method as will enable teachers and supervisors more easily and effectively to practice the concepts of the activity program." Furthermore, from the standpoint of character education, the same report presents the following conclusions:

1. Pupils in the activity program excel "in the development of civic attitudes and understanding of social relationships."
2. They like school better, find it more interesting, tend to carry its influence into life outside the school.
3. They tend to excel in such qualities as cooperativeness, self-confidence or poise, lack of subservience, creative abilities, self-discipline and scientific outlook.

F. *The Influence of Organization, Administration and Supervision on Character Education:*

1. *Philosophy Underlying the Structure, Administration and Supervision of American Education:*

a. Organization and Administration: A coherent, democratic philosophy of education requires that every aspect of the educational enterprise and every problem concerning its organization and operation be viewed in the light of the functions and purposes that American democratic education are intended to serve. The entire school must function in ways through which all engaged in its work may learn to live well together. The school must organize and administer its program "so as to make desirable conduct the most likely outcome" of training, thus contributing to the making of wholesome personality. However, the school cannot accomplish this single-handed. Active cooperation of parents and public is necessary in those matters in which they have a common responsibility. The foundations of many desirable as well as undesirable habits, attitudes, and dispositions, are laid in the home, on the street, and in other places outside the school. All agencies affecting the child must consciously and systematically work toward the development of good character in children. To insure effective cooperation of these agencies with the school, it is important that the school keep the parents and the public informed regarding its aims, objectives, and program. The National Education Association has performed exceptional service in this respect through planning the annual observance of American Education Week, for which it prepares and publishes excellent material. The New York City educational system in cooperation with the National Education Association, the United Parents' Association and other organizations provides for the observance of Open School Week. In addition, the United Parents' Association, the Public Education Association, the Teachers' Guild Associates, the local Parents' Associations and Mothers' Clubs do excellent work in studying the operation of the schools, disseminating reliable

information about them, and working for their improvement.

These activities, together with the presentation of pageants, plays, commencement exercises, exhibits of pupils' work, the conduct of philanthropic, welfare and civic enterprises conducted through the cooperation of the school and community; student publications, radio talks, contributions to daily papers; special emphasis upon the behavior of pupils in public places, such as street cars, buses, subways—all serve to bring about a closer association between the school, the parents, and the public, making for better character training.

The keynotes of the philosophy underlying the democratic organization, administration, and supervision of the school system for improved character training are:

a. Regard for the worth of each person.
b. Responsible freedom.
c. Intelligent cooperation.
d. Concern for the general welfare.
e. Reason, consultation and persuasion.

The spirit and the attitudes involved in this philosophy should prevail in all the relationships involved in the educational enterprise. This frame of reference may be presented schematically as follows:

RELATIONSHIPS	PHILOSOPHY OF ORGANIZATION AND ADMINISTRATION				
	Regard for the worth of each person	Responsible freedom	Intelligent cooperation	Concern for the general welfare	Reason, consultation and persuasion
Public					
School Board					
Associate Superintendents					
Assistant Superintendents					
Principals					
Teachers					
Pupils					

The application of such a philosophy should be directed toward the achievement of a program of organization and administration that will provide:

a. A high type of leadership among administrators, supervisors, principals and teachers.
b. A school environment and facilities that are conducive to the achievement of the best results.
c. A vital curriculum attuned to the needs of growing individuals and a changing culture.
d. Socialized curricular and co-curricular activities.
e. Democratic participation in the life of the school by all who are concerned from the public down to the individual child.

The organization and administration of the school on a democratic basis is of such importance for character training that a more detailed consideration of this point is necessary. It can be promoted by general, departmental, grade and individual conferences in which an interchange of ideas among supervisors and teachers is encouraged; by investigations, surveys, reports and discussions of vital educational problems; and by a functional organization of the staff into committees for studying various problems and performing different tasks. The committees should be organized democratically on the basis of voluntary choice; effectively, on the basis of interests and capabilities and clearly defined functions. Their findings and recommendations should be reported, discussed, appraised and acted upon. The writer presents below a suggestive list of committees which he has tried out with reasonable success in his school.

SCHOOL COMMITTEES AND THEIR FUNCTIONS

1. Club Activities—to plan, administer and supervise the activities of special interest clubs.
2. Methods of Instruction—to study and report on the best methods of teaching and of learning suitable for our pupils.
3. School Publications—to plan, administer and supervise the publication of school and class bulletins, newspapers and magazines.

4. School Exhibitions and Entertainments—to plan, arrange and supervise school and classroom exhibitions of pupils' work and activities, assembly programs and school entertainments.
5. School Decoration—to plan and supervise the decorations of classrooms and other parts of the school building.
6. Tests and Measurements—to plan, administer, and supervise the administration of scientific tests and measurements of pupils' intelligence, scholastic attainments, character traits, behavior problems, etc.
7. Preventive, Diagnostic and Remedial Teaching—to study and report on the best means and methods for the prevention and correction of scholastic difficulties in reading, arithmetic and other subjects.
8. Visual, Auditory and Objective Aids—to plan and supervise the use of visual auditory and objective aids to education.
9. Attendance—to study the attendance problems of our school in order to determine causes of non-attendance with a view to recommending appropriate measures for the improvement of attendance.
10. Curriculum—to study and report on the problems of the modification and adaptation of the curriculum to the individual needs, capacities and interests of our pupils.
11. Child Welfare—to study and provide for the health, safety, recreational and related needs of our children, so as to promote their physical, mental and moral well-being.
12. Child Guidance—to study and plan a program of activities designed to develop desirable habits, attitudes and traits of good character, and to prevent and correct undesirable ones.
13. Social Committee—to plan and arrange social activities for the staff, and to provide other suitable means and opportunities for the promotion of a spirit of friendliness, sociability and mutual respect and understanding.
14. Teachers' Welfare—to consider and recommend intelligent, practicable measures pertaining to problems arising out of the organization, administration and supervision of our school, for the promotion of teacher welfare, morale and esprit-de-corps, in conformity with the by-laws of the Board of Education, the rules, regulations and orders of administrative and supervisory officials, and the highest standards of social and professional ethics.

The importance of democratic principles and their implications for educational practice have been the subject of intensive study during recent years. The series of reports of the Educational Policies Commission of the National Education Association, (15, 16, 17, 18, 19, 20) the proceedings of the Administrative Officers of Public and Private Schools, held by the Department of Education of the University of Chicago, (32) the work of Professor Thomas H. Briggs, and the pending survey of the Committee on Practical Democracy in Education of the Board of Education of the City of New York, all deal with this problem. In the last named inquiry eighteen of the most important characteristics of democracy pertaining to education are set forth and the heads of schools, in cooperation with their teachers, are asked to study and report how each characteristic is notably exemplified, notably ignored or denied and how it can be better exemplified, in connection with "all of the interrelations between pupils, teachers, and administrators, parents and the community, and also the choice and use of organization and administration, curricula, methods of teaching, direction of extra-curricular activities, the social life of the school and discipline."

b. *Supervision.*

Of the utmost importance for the improvement of general education and character education in the elementary schools of the City of New York is the acceptance of a more comprehensive conception of the nature of supervision in harmony with the scientific and philosophic points of view emerging in contemporary educational theory. At two points in the report on the recent survey of supervision in the elementary schools of the City of New York, the following definitions of supervision occur.

> "Supervision sets the tone and establishes the philosophy of the school. It deals mainly with the improvement of teachers and the teaching act. It has two phases, actual classroom observation to discover needs and a program required to meet these needs." (4: p. 22.)

"It is generally acknowledged that supervision is concerned with such matters as the interpretation of the curriculum, the improvement of the teaching act, the improvement of teachers in service, the measurement of results, the investigation of educational problems, the building of morale, and the development of public relations." (4: p. 28.)

Let us examine these definitions in the light of recent changes in the concept of supervision. The first modern statement formulated by Burton in 1922 includes: (2: Chap. I.)

(1) Improvement of the teaching process.
(2) Improvement of teachers in service.
(3) Selection and organization of subject matter.
(4) Testing and measuring.
(5) Rating of teachers.

In 1930, he presented a summary statement based upon objective studies and theoretical formulations:

(1) Studying the Teaching-Learning Situation
(a) The aims and objectives of education and of supervision
(b) The pre-requisites of satisfactory pupil development and achievement
(c) The capacities, interests, attitudes, and habits of the pupils
(d) The teaching
(e) The learning
(f) The curriculum
(g) Instructional materials and equipment
(h) The outcomes
(i) Social environment (community)

(2) Improving the Teaching-Learning Situation

(1) *Inspection*, including survey of entire school system, materials of instruction, equipment, teaching staff, pupils, teaching, learning, etc.
(2) *Training*, by means of
  (a) Direct improvement of teaching.
  (b) Improvement of teachers in service.
(3) *Guidance*, through
  (a) Organizing programs of cooperative activity.
  (b) Developing and maintaining good professional spirit.
(4) *Research*.
(5) *Professional and semi-administrative duties*.

Within the past three or four years Barr, Burton and Brueckner have called attention to the emerging concept of supervision which Brueckner calls "organismic supervision," (7: pp. 435–441.) revealing certain definite trends:

(1) Increasing objectivity (scientific).
(2) Emphasis upon democratic philosophy and policy.
(3) More comprehensive view.
(4) Distribution of supervisory duties among a larger number of persons.

"Organismic supervision" has three major functions and a number of related minor ones. They are—

(1) Studying the teaching-learning situation.
(2) Improving the teaching-learning situation.
(3) Evaluating the means, methods, and outcomes of supervision.

The first two functions, which are receiving most attention at the present time, may be presented schematically as follows:

The third major function involves the analysis and appraisal of supervisory procedures with a view to their improvement and includes such problems as:

(1) The development and application of techniques for appraising supervision.
(2) Appraising the general value of supervision.
(3) Appraising the relative value of various supervisory plans.
(4) Appraising and improving the personnel of supervision.

Such a comprehensive, coherent conception of supervision, as part of an equally comprehensive and coherent philosophy of general education and of character education, as herein above described, operating in the

elementary schools of the City of New York must, of necessity, exercise a profound influence on the organization and administration of character training, for it involves no less than organizing the whole educational situation so that parents and public may cooperate with administrators, supervisors, principals, teachers and pupils to act, each in his sphere, in accordance with the fundamental standard of moral responsibility—"integration of values—doing the best thing in each situation."

The conception of character education expounded in this article can be summarized in the following schematic outline, which is self-explanatory. It is founded upon an organismic, purposive and integrative view of social development and individual development, with democratic education as the chief means and moral responsibility as the standard of values.

Bibliography

1. *Annual Report of the Assistant Superintendents of Schools*, City of New York, School Year 1937–1938. New York: Board of Education, Revised as of 1938–1939.
2. Barr, A. S., Burton, W. H., and Brueckner, L. J. *Supervision.* New York: D. Appleton–Century Company, 1938.
3. Board of Education, City of New York. *Report on the Activity Program in the Schools of the City of New York.* New York: Board of Education, 1941.
4. Board of Education, City of New York. *Supervision in the Elementary Schools of the City of New York.* Report of the Committee on Supervision Appointed January, 1938, to carry out Recommendation 44, Committee on Maladjustment and Delinquency. (Mimeographed) New York: Board of Education, 1941.
5. Borgeson, F. C. *All School Activities.* New York: A. S. Barnes and Company, 1931.
6. ------ *Group Interest Activities.* New York: A. S. Barnes and Company, 1931.
7. Brueckner, L. J., and Koenker, Robert. "Organismic Supervision," *Elementary School Journal,* XL (February, 1940), 435–441.
8. *Cardinal Principles of Secondary Education.* Bulletin No. 35. Washington, D. C.: Bureau of Education, 1918.

Organic and Synoptic View of Character Education

	Social and Educational Philosophy	
1. Aims and 2. Objectives derived from social and educational philosophy	Major Aims of Education Social Self-Realization Individual Self-Realization | Competency in Social Relationships | Economic Efficiency and Well Being | Civic Responsibility Subsidiary Aims of Education Educational Objectives in terms of understandings, abilities, skills, habits, attitudes, ideals, and dispositions required by the individual to be a wholesome person and to live the good life in a progressive society which he helps to improve.	Standard of Values *Moral Responsibility*—"integration of values—doing the best thing in each situation."
3. Means Curriculum derived from culture, the accumulated experience of the race	The experiences pupils shall engage in to promote their growth and development toward the attainment of the individual and the social goals of education Experience Curriculum Based upon social functions and areas of living Sequence of "units of conduct" or "units of life" Each unit planned in terms of objectives, together with materials and activities required to realize them	
4. Methods a. Teaching & b. Learning c. Organizing d. Administering & e. Supervising	Functional and integrative, individualizing and socializing Functional, integrative, scientific, creative, cooperative and democratic	
5. Results Outcomes and their Appraisal	In terms of behavior in real life situations and effect upon growth and development in self-realization, both individual and social	

9. Caswell, H. L., and Campbell, D. S. *Curriculum Development.* New York: American Book Company, 1935.
10. Department of Superintendence Tenth Yearbook: *Character Education.* Washington, D. C.: Department of Superintendence, National Education Association, 1932.
11. Dewey, J. *Democracy and Education.* New York: Macmillan Company, 1916.
12. ------ *Experience and Education.* New York: Macmillan Company, 1916.
13. ------ "Some Aspects of Modern Education," *School and Society,* XXXIV (October, 1931), 579–584.
14. Educational Policies Commission. *Education and Economic Well-Being in American Democracy.* Washington, D. C.: Educational Policies Commission, National Education Association, 1940.
15. ------ *Learning the Ways of Democracy:* A Case Book in Civic Education. Washington, D. C.: Educational Policies Commission, National Education Association, 1940.
16. ------ *Social Services and the Schools.* Washington, D. C.: Educational Policies Commission, National Education Association, 1939.
17. ------ *The Education of Free Men in American Democracy.* Washington, D. C.: Educational Policies Commission, National Education Association, 1941.
18. ------ *The Purposes of Education in American Democracy.* Washington, D.C.: Educational Policies Commission, National Education Association, 1938.
19. ------ *The Structure and Administration of Education in American Democracy.* Washington, D. C.: Educational Policies Commission, National Education Association, 1938.
20. ------ *The Unique Function of Education in American Democracy.* Washington, D. C.: Educational Policies Commission, National Education Association, 1937.
21. Fite, W. *Moral Philosophy: The Critical View of Life.* New York: The Dial Press, 1925.
22. Harap, H., et al. *The Changing Curriculum.* New York: Appleton–Century Company, 1937.
23. Hopkins, L. T., et al. *Integration—Its Meaning and Application.* New York: Appleton–Century Company, 1937.
24. Jordon, R. H. *Extra-Classroom Activities in Elementary and Secondary Schools.* New York: Thomas Y. Crowell Company, 1928.
25. Kefauver, G. N. *An Educational Platform for 1936.* Stanford University, California: Stanford University Press, 1936.
26. McKown, H. C. *Activities in the Elementary School.* New York: McGraw–Hill Book Company, Inc., 1938.
27. Melvin, A. G. *The Technique of Progressive Teaching.* New York: John Day Company, 1934.
28. National Society for the Study of Education. Twenty-fifth Yearbook: Part II. *Extra-Curricular Activities.* Bloomington, Illinois: Public School Publishing Company, 1926.
29. National Education Association. *Implications of Social Economic Goals for Education.* Washington, D. C.: National Education Association, 1937.
30. ------ "Education for Character," *Research Bulletin,* XII, Nos. 2 and 3. Washington, D. C.: National Education Association, 1934.
31. ------ *The Classroom Teacher and Character Education.* Seventh Yearbook: Department of Classroom Teachers. Washington, D. C.: National Education Association, 1932.
32. Reavis, W. C. *Democratic Practices in School Administration.* Chicago: University of Chicago Press, 1939.
33. Stratemeyer, F. *The Effective Use of Curriculum Materials.* Contributions to Education, No. 460. New York: Bureau of Publications, Teachers College, Columbia University, 1931.
34. White, A. K., and Macbeath, M. A. *The Moral Self, Its Nature and Development.* New York: Longmans, Green and Company, 1923.

OBJECTIVITY IN THE EVALUATION OF PERSONALITY

T. W. Richards and Willis Ellington
The Samuel S. Fels Research Institute
Antioch College, Yellow Springs, Ohio

When the psychologist sits down to the task of evaluating the other fellow's personality, he does so knowing that he may well inject some of himself into his appraisal. This source of error has furnished the incentive for standardization of clinical training, on the theory presumably that everyone should see things the same way, even if there might develop a common bias. Experimentalists, on the other hand, have been eager to refine methods of appraisal, to reduce the personal equation. One often suspects that the premium placed on objectivity has sometimes sacrificed importance to accuracy.

Somewhere in the ground of compromise between clinical significance and scientific accuracy the rating scale technique appears as an approximation to both. In the catalogue of psychometrics this unhappy position of compromise gives the rating scale an uncertain reputation,—a sort of quasi-scientific standing which variably satisfies the odd assortment of people who seem to find it useful. Yet in this very technique itself, it seems to us, there is an area of real psychological importance. Possibly an analysis of the rating function might yield some insight into the personal equation. You may call it bias, if you will, or the validity of opinion, or—grossly—the psychology of rating. We have set ourselves the task here, then, of examining the product of the rater. The basic technique used is the analysis of intercorrelations among the ratings made by the single rater on several personality characteristics.

In making a series of ratings the rater draws on the sum-total of his experience. This experience may be of two sorts:

(1) The experience the rater has with the group to be rated. This represents his sampling of the population.

(2) The extra-sampling experience of the rater. This includes such things as his familiarity with psychological variables, with the rating task, and with the particular rating scales to be used. Important here also is his functional psychology itself; that is, the psychological frame of reference which aids in his thinking when he evaluates personality.

The "true behavior" of the rated population, to the extent that it impresses itself upon the rater, manifests itself through the sampling experience.

Any series of ratings, then, represents expression of an alloy of experience which the rater has had.

The ratings used in this report were made on school children in the elementary grades by school teachers.[1] The traits rated were defined[2] for the teachers as follows:

1. Good Behavior—Misbehavior: Consider the child's conduct, his conformity to the classroom program. Is he a typically "good" child, behaving as he is supposed to and needing no discipline? Or does he continually fail to accept authority, making trouble and requiring discipline?
2. Interest in School Work: How much interest does the child show in his school work? Is he eager to learn, easily stimulated by new work? Or does he dislike school work, find it distasteful, uninteresting, or even hate it?
3. Energy Output: Consider the child's energy output. Does this child put great energy into every activity; does he seem to be power-driven? Or does he try to get by with the least possible exertion?
4. Endurance: Consider the effect of physical exertion on the child. Does the child tire quickly even after the slightest exertion? Or does he have a great deal of endurance, working or playing hard for long periods of time without getting tired?
5. Emotionality: How often does the child react emotionally? Does he very frequently become excited, angry, afraid, or

[1] It is very important throughout the remainder of this paper to keep in mind the fact that the *raters* were also *teachers*. This means that in comparing the ratings of one rater with those of another, the personal factor of the rater could enter in both by affecting the ratings directly and by affecting the "true behavior" of the ratees.

[2] These definitions and the scales themselves are to be published together with data on their reliability in a forthcoming publication of the Fels Research Institute.

disappointed; does he often have temper tantrums, sob or cry; does he often worry, sulk, become antagonistic, negativistic? Or, on the other hand, is he apathetic, rarely showing any emotion?

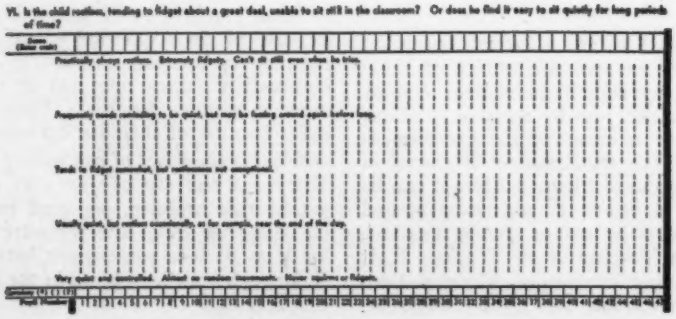
6. Restlessness: Is the child restless, tending to fidget about a great deal, unable to sit still in the class-room? or does he find it easy to sit quietly for long periods of time?
7. Sensitiveness: Consider the child's concern for the opinion of others, his desire to be approved. Is he extremely sensitive to what others think of him? Or is he indifferent, oblivious to group opinion? (The sensitive child may show acute embarrassment, or he may strut and pose with great self-assurance).
8. Dominance: Consider the child's relation to his classmates and other pupils his own age. Does he try to dominate others and run the show, resisting rival initiative? Or does he accept domination, aggression, suggestions, uncritically and without resistance, when coming from other children?
9. Lack of Emotional Control: No matter how seldom or often the child shows emotion, what does he do when he is emotional? Does he stifle his grief into quiet sobbing, never crying openly; does he try not to show it when he is angry or afraid? Or does he lack emotional inhibitions, never trying to hide his feelings?
10. Concentration—Distractibility: Consider the child's habits of attention to the main task in work and play. Is he so easily distracted that he can't pay attention for more than a moment? Or is he such a good concentrator that once he gets involved in something it is almost impossible to get him away from it?
11. Impulsiveness: Consider the child's tendency to control his impulses of the moment and to act only with deliberate consideration. Is he inclined to be deliberate, indecisive? Or is he impulsive, headlong, impetuous, precipitate, impatient?
12. Neatness—Carelessness: How precise is the child's work? Does he do his work carefully, painstakingly, insisting upon exactness? Or does he work haphazardly, consistently trying to get by with as little pains as possible, content with careless work?

A sample of the scale format, showing the nature of the cues, appears on this page. This is the scale for *restlessness* 6.

The intercorrelations for one rater (Teacher A) appear in the lower portion of Table I. An analysis[3] of this table (by means of Thurstone's centroid method of factor analysis) showed a two-factor pattern for this teacher. As shown in the upper left of Figure I it will be seen that with *energy* 3, and *endurance* 4 at the opposite end of (non) *interest in school work* 2 we have a factor of PARTICIPATION,[4] orthogonal to another factor, the most axial scales of which are *restlessness* 6 and *misbehavior* 1. This second factor we have called BAD BEHAVIOR. Note that (non) *concentration* 10 and (non) *neatness* 12 lie between BAD BEHAVIOR and the negative end of the PARTICIPATION axis, while between BAD BEHAVIOR and PAR-

[3] The authors wish to acknowledge the assistance of Rita Senf in making several of the statistical analyses here involved.

[4] The authors do not insist on the names they have given to the factors revealed in these analyses. The reader is free to name them as he will.


VI. Is the child restless, tending to fidget about a great deal, unable to sit still in the classroom? Or does he find it easy to sit quietly for long periods of time?
Some
(Some one)
Practically always restless. Extremely fidgety. Can't sit still even when he tries.
Frequently needs something to be quiet, but may be holding control again before long.
Tends to fidget somewhat, but restlessness not exceptional.
Usually quiet, but restless occasionally, as for example, near the end of the day.
Very quiet and controlled. Almost no noticeable improvements. Never squirms or fidgets.

TICIPATION are the emotional variables (5, 9), *impulsiveness* 11, and *dominance* 8. *Sensitiveness* 7 has negligible loadings on both factors.

How well does the pattern of one rater agree with that of another? We performed the same type of analysis on the intercorrelations between the ratings of Teacher T. Teacher T was a student of education, who did practice teaching in a progressive private school. She was considerably younger and less experienced than Teacher A. She rated an entirely different group of children. Intercorrelations for this teacher's ratings are shown in the upper half of Table I.

TABLE I

INTERSCALE CORRELATIONS BETWEEN THE RATINGS OF TEACHER A (LOWER HALF) AND TEACHER T (UPPER HALF)

Teacher A	1	2	3	4	5	6	7	8	9	10	11	12	Teacher T
1		—.65	+.23	+.02	+.53	+.78	+.64	+.31	—.85	+.76	—.71	—.84	
2	—.13		+.14	—.21	—.34	—.54	—.52	—.06	+.72	—.53	+.39	+.62	
3	—.06	+.72		—.78	—.01	—.21	+.08	+.78	+.02	—.23	+.03	+.23	
4	—.09	—.70	—.88		+.41	+.49	+.31	—.74	—.27	+.49	—.34	—.51	
5	+.60	+.29	+.30	—.41		+.71	+.64	—.10	—.59	+.54	—.63	—.66	
6	+.79	+.13	+.16	—.31	+.53		+.81	—.20	—.67	+.87	—.76	—.89	
7	—.03	+.33	+.19	—.12	+.11	+.04		—.12	—.45	+.67	—.68	—.71	
8	+.55	+.55	+.59	—.74	+.70	+.60	+.04		—.23	—.18	+.13	+.17	
9	—.18	—.25	—.16	+.15	—.37	—.46	—.26	—.26		—.66	+.61	+.79	
10	+.62	—.40	—.51	+.32	+.30	+.61	—.17	+.03	—.30		—.84	—.86	
11	—.41	—.23	—.13	+.19	—.61	—.28	—.24	—.33	+.25	—.08		+.76	
12	—.47	+.49	+.69	—.52	+.08	—.29	+.08	+.24	—.14	—.60	+.16		

The pattern for Teacher T also provided two factors. This pattern is shown in the upper right of Figure I. One of her two factors, composed of *energy* 3, *dominance* 8 and *endurance* 4 we have called an ACTIVENESS factor. Orthogonal to this was a factor of BAD BEHAVIOR, composed of *misbehavior* 1, the emotional variables (5, 9), *sensitiveness* 7, *impulsiveness* 11, (non) *interest in school work* 2, (non) *neatness* 12, *restlessness* 6 and (non) *concentration* 10.

In an attempt to get a better conception of the differences and similarities of these two patterns, we superimposed one of these plots on the other, and by suitable rotations tried to equate as many of the variables as was possible. The final product is shown in the lower diagram of Figure I. Here it can be seen that *misbehavior* 1, *endurance* 4, and (non) *concentration* 10 are almost identical. Moving from Teacher A to Teacher T, it can be seen that the rather diffuse pattern of Miss A moves into two tight clusters for Miss T. The major shift is shown by (non) *interest in school work* 2, which goes from the negative end of the PARTICIPATION axis for Teacher A to the BAD BEHAVIOR axis of Miss T. The emotional variables (5, 9) and *impulsiveness* 11, show this shift to a lesser extent from the positive end of the PARTICIPATION axis, while *sensitiveness* 7 goes from a negligible position for Miss A to the BAD BEHAVIOR axis for Miss T. This "gathering together" takes place to a lesser extent in the other factor, with *dominance* 8 and *energy* 3 making shifts to a more central position on the PARTICIPATION-ACTIVENESS axis. It seems as if the psychological variables each teacher rated assumed a more variable pattern of correlation with the two main factors for Teacher A than they did for Teacher T. The latter seems to have made her ratings with reference to a two factor pattern much more closely than did Teacher A; or stated another way, the two factor pattern for Miss T formed a much more rigid frame of reference than it did for the other rater.

These teachers could give different results because of their differing sampling of child behavior or because of differing adequacy of observation (since one was a regular, the other a practice teacher). They might also differ because of their background.

What is the effect of controlling the population rated and the adequacy of sample? To get at this problem, we used ratings made on Miss A's group the semester previously, by Miss K. The correlations between A and K for each scale are shown in the first column of Table II.

TABLE II

INTERMEMBER CORRELATIONS FOR THREE PAIRS OF RATERS

	Pairs of Teachers A&K	T&M	P&S
NUMBER OF CASES	31	19	43
1. Misbehavior	.51	.58	.81
2. Interest in school work	.47	.71	---
3. Energy	.60	.37	.67
4. Endurance	.47	.65	.63
5. Emotionality	.03	.12	.67
6. Restlessness	.64	.15	.84
7. Sensitiveness	.05	—.24	.76
8. Dominance	.51	.25	.75
9. Lack of emotional control	.18	.23	.69
10. Concentration	.79	.35	.66
11. Impulsiveness	.18	.16	.64
12. Neatness in school work	.61	.25	.78
Root mean square	.48	.38	.72

The patterns shown by Miss A and Miss K are compared in Figure II. The factor pattern presented by Teacher K was in some respects very similar to that shown by Teacher A. Again the PARTICIPATION factor appeared, and orthogonal to it was *restlessness* 6. But the GOOD-BAD BEHAVIOR continuum was less clear. When both factor patterns are presented in one figure, with the PARTICIPATION cluster superimposed, a very interesting relationship is demonstrated: Teacher K (the square plots) rated *impulsiveness* 11 and to some extent the emotional variables (5, 9) near the PARTICIPATION cluster. *Concentration* 10 and *neatness* 12 were also very close to PARTICIPATION. *Sensitiveness* 7 was very close to (non) *misbehavior* 1. Teacher A considered the emotional variables (5, 9) and *impulsiveness* 11 to be very close to *restlessness* 6, and all this was at the opposite pole to GOOD BEHAVIOR. Thus this restless, emotional, impulsive picture defines BAD BEHAVIOR for her. *Concentration* 10, moves slightly away from PARTICIPATION toward GOOD BEHAVIOR. *Sensitiveness* 7, though it is not BAD BEHAVIOR, certainly loses the GOOD BEHAVIOR meaning given to it by Teacher K. Reference to Table II shows that the scales which in Figure II shift most when we go from Teacher K to Teacher A have the smallest coefficients of correlation between raters. For the two emotional scales (5, 9), for *impulsiveness* 11 and for *sensitiveness* 7 the coefficients are uniformly below .20.

How much more alike are the patterns of the teachers who rated the same group than the patterns of teachers who rated different groups? In other words, is Teacher A more like Teacher K than she is like Teacher T? Comparison of Figures I and II suggests that she is. In other words, sampling differences may be a factor in producing different teacher patterns. But essentially there are almost as radical shifts in the same-group comparison as in the different-group comparison. That is, even under conditions of considerably more similar sampling, raters differ.

What is the effect of similarity of background? To answer this question we com-

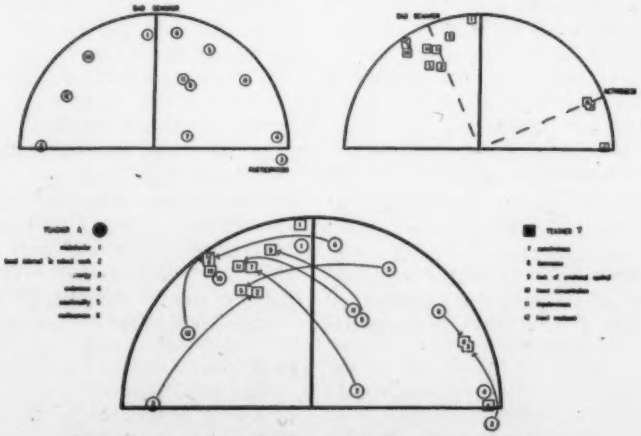


Figure I.—Factor Patterns for Teachers A and T, Singly and Superimposed.

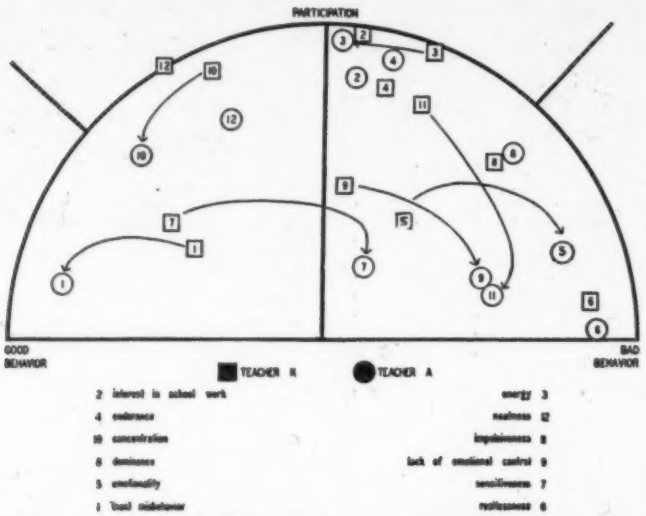


Figure II.—Factor Patterns for Teachers A and K Superimposed.

pared three pairs of teachers, each pair rating the same children, the pairs varying in the degree of similarity in their background. Teachers A and K were two public school teachers. Although differing in age and experience, they each possessed the schoolteacher's training. We wanted to compare other pairs of raters, just as we had compared Teachers A and K, with the similarity in training varied. Two other pairs were used. The first pair, Teachers M and T (compared above with Teacher A) were widely divergent in their backgrounds. One, Teacher M, was a man, a teacher in a progressive private school, father of several children. Teacher T was a woman, a student teacher of the group taught regularly by Teacher M. She had a slightly different sample of the children's behavior because of her role as occasional observer and practice-teacher, but the raters did their ratings over approximately the same period.

In the case of the second pair of raters (T & M), comparison graphically was difficult, since one rater (T) showed two factors while the other (M) showed three. Comparing the patterns indicated that each of the raters showed an ACTIVITY factor of *energy* 3, and *endurance* 4. Whereas the first rater (T, the student) tended to lump the remaining scales into a factor of BAD BEHAVIOR (the less desirable end of the scales having positive loadings) the regular teacher (M) split this BAD BEHAVIOR factor into two. One of these components for teacher M tended toward *misbehavior* 1, and included *emotionality* 5, *restlessness* 6, *impulsiveness* 11, and *lack of emotional control* 9, while the other factor, orthogonal to the emotional misbehavior was made up of the traits of poor school work: (non) *interest in school work* 2, (non) *concentration* 10, (non) *neatness* 12.

In Table II we see that this pair of raters showed the least agreement between them of all the pairs. Of the inter-rater correlations for these two teachers, only five are above .30: those for *interest in school work, energy, endurance* (which form the ACTIVITY factor on which they agreed), and (non) *misbehavior* and *concentration.*

These raters, Teachers M and T, were less alike in their patterns than were Teachers A and K. Indeed, Teacher T showed less similarity to Teacher M, with whom she rated the same group of children than she did with Teacher A or Teacher K, each of whom rated a different group than did Teacher T. Like A and K, T showed two factors.

The third pair of raters were persons of relatively similar training. They were staff members of the Institute, one the teacher of the group, the other the psychological ob-

server. Each had had graduate training in child psychology in the same University department. They were of like age. They discussed the children in their groups freely and with mutual understanding. Of course, their ratings were made independently. Inter-rater correlations between the ratings for each pair appear in Table II. Note that of the three pairs of raters, P & S are in closest agreement.

Each of the third pair of raters presented a three-factor pattern. The relationships between their patterns are shown in Figure III. This is a schematic diagram with the points plotted as accurately as possible on a two-dimensional projection of a three-dimensional figure. The factor loadings are presented in Table IV.

Teacher S (the circles) showed the three factors in very clear distinction. One was formed alone by *sensitiveness* 7 (upper left), another by the ACTIVITY variables: *energy* 3 and *endurance* 4. (These teachers did not rate *interest in school work* 2 since theirs was not a classroom situation.) The third factor was composed of (non) *concentration* 10 and (non) *neatness* 12. The pattern of the second rater (P) is basically in very close agreement with that of Teacher S. Certain shifts are apparent, however. The emotional variables (5, 9) shift from a multifactor posi-

TABLE III

INTERSCALE CORRELATIONS BETWEEN THE RATINGS OF TEACHERS M, K, P & S

Scale	1	2	3	4	5	6	7	8	9	10	11	12	
1		−.28	−.18	+.01	−.11	+.39	−.19	+.06	+.27	+.35	−.02	−.41	
2	−.16		+.89	−.83	+.35	+.22	+.30	+.66	−.43	−.78	−.78	+.79	
3	−.30	+.41		−.79	+.42	+.45	+.13	+.69	−.48	−.66	−.72	+.68	
4	+.05	−.15	−.60		−.17	−.27	−.13	−.56	+.34	+.72	+.70	−.70	Teacher K
5	+.65	+.17	−.35	+.14		+.44	+.17	+.34	−.28	−.10	−.15	+.17	
6	+.75	+.01	−.31	+.15	+.84		−.35	+.47	−.17	+.04	−.26	−.39	
7	+.06	+.24	−.34	+.28	+.15	−.05		−.10	−.15	−.46	−.15	+.56	
8	+.23	+.21	+.33	−.22	−.03	−.14	+.09		−.06	−.25	−.72	+.28	
9	−.66	+.13	+.20	+.13	−.70	−.71	−.06	+.03		+.30	+.35	−.33	
10	+.65	−.24	−.23	+.17	+.23	+.33	+.10	−.21	−.21		+.54	−.92	
11	−.70	+.21	+.36	−.30	−.64	−.66	−.30	+.02	+.54	−.55		−.56	
12	−.59	+.52	+.49	−.26	−.13	−.23	−.29	−.29	+.13	−.77	+.51		
1			+.58	−.11	+.68	+.76	+.16	+.55	−.47	+.51	−.60	−.54	
2													
3	+.57			−.65	+.60	+.76	+.10	+.45	−.47	+.24	−.72	−.18	
4	−.56		−.89		−.17	−.29	+.17	−.27	+.09	+.23	+.32	−.12	Teacher S
5	+.59		+.34	−.22		+.75	+.40	+.67	−.70	+.46	−.56	−.47	
6	+.71		+.75	−.61	+.64		−.06	+.57	−.60	+.47	−.78	−.50	
7	+.29		+.05	+.07	+.56	+.16		+.39	−.10	+.21	−.05	−.03	
8	+.66		+.45	−.41	+.57	+.48	+.29		−.51	+.09	−.55	−.12	
9	−.60		−.23	+.17	−.89	−.59	−.47	−.55		−.48	+.57	+.56	
10	+.55		+.39	−.30	+.45	+.66	+.11	+.28	−.57		−.44	−.84	
11	−.66		−.78	+.71	−.54	−.83	−.18	−.49	+.50	−.67		+.43	
12	−.62		−.52	+.35	−.53	−.66	−.43	−.39	+.56	−.70	+.59		

Teacher M / Teacher P

TABLE IV

FACTOR LOADINGS FOR EACH SCALE FOR EACH OF TWO TEACHERS (P & S)

	Teacher P				Teacher S			
	K_1	K_2	K_3	h^2	K_1	K_2	K_3	h^2
1	+.823	−.052	−.172	.710	+.793	+.075	−.074	.640
2								
3	+.738	+.513	−.319	.910	+.765	+.291	+.446	.869
4	−.634	−.471	+.462	.837	−.280	−.195	−.638	.523
5	+.781	−.478	+.169	.867	+.862	−.337	−.064	.861
6	+.871	+.250	+.099	.831	+.863	+.277	+.129	.839
7	+.382	−.465	+.158	.387	+.223	−.489	−.142	.309
8	+.658	−.262	−.336	.615	+.671	−.436	+.323	.744
9	−.759	+.467	−.255	.859	−.727	+.118	+.094	.552
10	+.679	+.179	+.397	.651	+.605	+.185	−.657	.832
11	−.854	−.323	−.031	.834	−.805	−.185	−.197	.721
12	−.762	−.059	−.274	.674	−.607	−.239	+.635	.828

tion for rater S toward *sensitiveness* 7 for rater P, defining the sensitiveness factor for her as more clearly one of emotionality, less of ACTIVENESS. Both *concentration* 10 and (non) *neatness* 12 shift from their axial position taking on some of the character of emotionality. There are other minor shifts for 1, 3, 6 and 11.

These comparisons between teachers rating the same groups suggest that:

(1) Raters differ in the configurations they present when making their ratings.
(2) These differences exist even when considerable control is made of the dissimilarity of sampling.
(3) Although these differences suggest that the rater's personality affects the pattern he demonstrates in his rating function, they do not rule out differences in sample of a more subtle nature, such as the teacher's own effect on the population rated, etc.

How consistent is the pattern for the individual rater? The foregoing material shows that factor analysis of the rater's performance gives a picture of some of the ways in which he differs from other raters. The investigators wondered whether a single rater might differ in the course of time—whether his configuration would shift. In order to get at this problem the investigators analyzed the intercorrelations of one teacher's ratings (Teacher A) made at three different times, the spring semester (2nd grade), the fall semester and the following spring semester (3rd grade). For the most part she had the same group of children: from the first to the second rating, 18 children, from the second to third ratings, 17. From spring to the following spring there was a common population of 16.

In Table V appear the intercorrelations between each of the sets of ratings. It is apparent that agreement was greater between consecutive than between interspersed ratings. It was greatest between the last two series of ratings, possibly as a result of experience with the ratings, or because these ratings were made in the same school year. Certain ratings, however, showed a greater correlation from spring to spring than between consecutive semesters. These were *energy, endurance,* and *emotionality.*

TABLE V

CORRELATIONS BETWEEN THREE RATINGS OF TEACHER A ON THREE SUCCESSIVE SCHOOL GROUPS

	GROUPS		
	I&II Spring to Fall	I&III Spring to Spring 2	II&III Fall to Spring 2
NUMBER OF CASES	18	16	17
1. Mis-behavior	.74	.51	.88
2. Interest in school work	.60	.58	.70
3. Energy	.82	.84	.73
4. Endurance	.66	.81	.59
5. Emotionality	.33	.57	.26
6. Restlessness	.54	.44	.81
7. Sensitiveness	.68	.47	.34
8. Dominance	.31	.47	.65
9. Lack of emotional control	.23	.50	.66
10. Concentration	.85	.81	.82
11. Impulsiveness	.79	.51	.75
12. Neatness in school work	.71	.52	.66
Root mean square	.64	.60	.68

The interscale correlations for the first ratings of Teacher A appear in Table I. The intercorrelations for her second and third ratings appear in Table VI.

TABLE VI

INTERSCALE CORRELATIONS FOR SECOND AND THIRD RATINGS OF TEACHER A

	1	2	3	4	5	6	7	8	9	10	11	12	
1		−.11	−.00	−.13	+.40	+.80	+.28	+.12	−.40	+.34	−.53	−.35	
2	−.14		+.76	−.55	−.32	−.14	+.10	+.86	+.59	−.79	−.31	+.87	
3	+.38	+.70		−.63	−.01	+.16	+.40	+.79	+.37	−.73	−.31	+.61	
4	−.26	−.63	−.79		+.33	−.22	−.20	−.76	−.35	+.35	+.12	−.38	
5	+.59	+.29	+.66	−.30		+.57	+.62	−.32	−.73	+.24	−.22	−.50	
6	+.83	−.01	+.39	−.17	+.56		+.36	+.07	−.60	+.44	−.58	−.47	
$A_3$7	+.41	+.09	+.26	−.20	+.47	+.53		+.26	−.33	−.19	−.21	−.13A_2	
8	+.45	+.40	+.46	−.60	+.26	+.28	+.03		+.46	−.61	−.31	+.66	
9	−.14	+.02	−.15	−.32	−.45	−.36	−.24	+.24		−.72	+.08	+.81	
10	+.16	−.83	−.48	+.53	−.05	+.27	+.19	−.48	−.30		+.09	−.88	
11	−.68	+.03	−.41	+.32	−.37	−.83	−.34	−.34	+.16	−.19		−.08	
12	−.20	+.82	+.60	−.61	+.16	−.13	+.07	+.26	+.06	−.67	+.06		

TABLE VII

FACTOR LOADINGS FOR THREE SUCCESSIVE RATINGS

Teacher A

Scale	Spring K_1	Spring K_2	Spring K_3	Spring h^2	Fall K_1	Fall K_2	Fall K_3	Fall h^2	Spring K_1	Spring K_2	Spring K_3	Spring h^2
1	+.522	—.667	+.308	.812	—.311	—.695	—.339	.695	+.594	+.566	+.415	.845
2	+.607	+.576	—.217	.747	+.904	—.234	—.104	.883	+.641	—.645	—.288	.910
3	+.648	+.692	+.171	.928	+.757	—.494	+.195	.855	+.886	—.202	—.136	.844
4	—.731	—.542	—.276	.904	—.587	+.409	+.217	.559	—.713	+.464	—.212	.768
5	+.787	—.274	+.058	.697	—.478	—.523	+.535	.788	+.704	+.344	—.302	.705
6	+.690	—.531	+.224	.808	—.351	—.854	—.215	.898	+.641	+.684	+.199	.919
7	+.241	+.095	—.461	.279	+.039	—.606	+.527	.647	+.438	+.402	—.165	.381
8	+.839	+.068	+.332	.819	+.821	—.474	—.202	.940	+.573	—.233	+.480	.612
9	—.508	+.173	+.246	.349	+.784	+.443	—.242	.870	—.144	—.459	+.492	.474
10	+.088	—.836	+.176	.738	—.900	—.068	—.306	.909	—.402	+.772	—.020	.758
11	—.501	+.239	+.287	.390	—.086	+.569	+.227	.383	—.586	—.471	—.360	.695
12	+.222	+.727	+.053	.580	+.916	+.186	—.078	.880	+.523	—.621	—.310	.756

The factor loadings[5] for the three analyses for Teacher A appear in Table VII. In Figure IV an attempt is made schematically to compare her patterns.

In a general way, Teacher A shows in her three patterns factors of

Good-Poor School Work: *Energy* 3, *interest in school work* 2, *concentration* 10, *endurance* 4, *neatness* 12 (upper right)

Good-Bad Behavior: *Misbehavior* 1, *restlessness* 6, *impulsiveness* 11 (upper left)

General Emotionality - Placidity: General emotionality, variously including *lack of emotional control* 9, *emotionality* 5, and *sensitiveness* 7 (lower)

There are radical shifts in these constellations, however. The first constellation, GOOD SCHOOL WORK, is the most constant; the teacher's concept of good school behavior embracing the picture of careful, concentrated,

[5] Because Teacher A presented clearly a three factor pattern for her second and third ratings, the analysis of her first ratings was carried to three factors, though for Figures I & II only the two factor pattern is used.

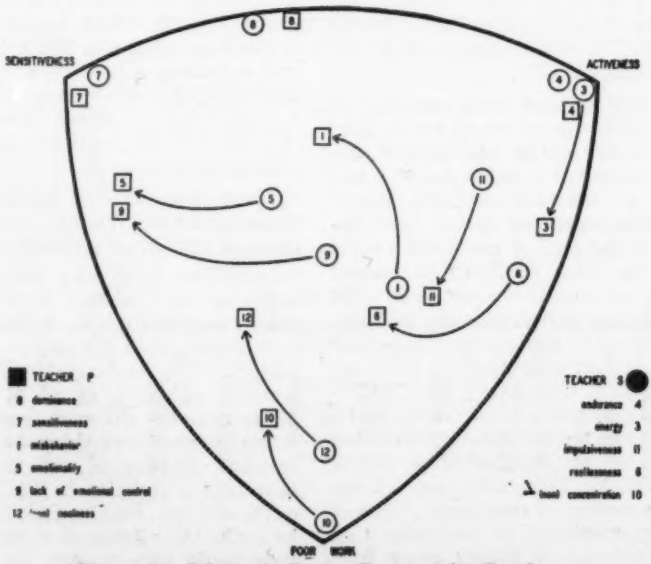


Figure 3.—Schematic Factor Patterns for Teachers P and S Superimposed.

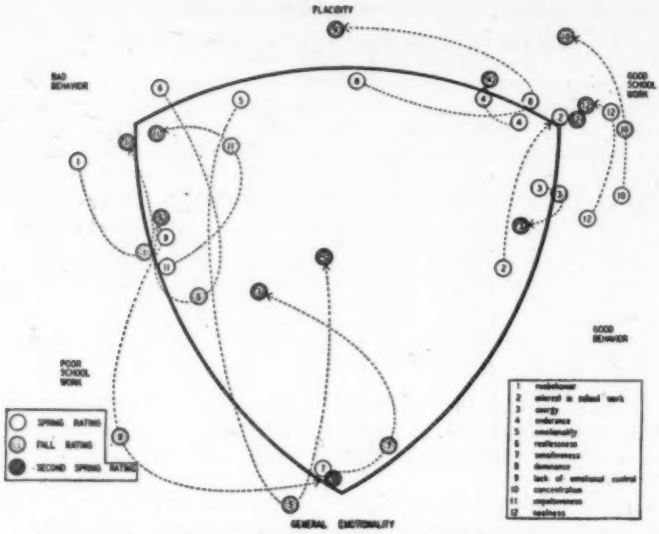


Figure IV.—Schematic Factor Patterns for Three Successive Ratings by Teacher A.

enthusiastic work is a fixed clear one. In her fall ratings, but not in her two spring ratings, this GOOD SCHOOL WORK factor included *dominance* 8. In the spring *dominance* 8 shared equally the BAD BEHAVIOR factor and in the second spring rating it also shared some PLACIDITY or the opposite of GENERAL EMOTIONALITY.

The next most constant factor was that of BAD BEHAVIOR (upper left in Figure IV). In the first rating BAD BEHAVIOR approximated *restlessness* 6, *misbehavior* 1, and *emotionality* 5. Close to it but fairly involving the general emotional factor were *impulsiveness* 11 and *lack of emotional control* 9. That is, the BAD BEHAVIOR pattern was definitely emotional—the emotional child was "bad." In the fall pattern the variables of emotionality (5, 9) shifted away from BAD BEHAVIOR toward *sensitiveness* 7 as though to define a pole of GENERAL EMOTIONALITY. Notice that *lack of emotional control* 9, not only moves toward *sensitiveness* 7, but it acquires a sizable significance of POOR SCHOOL WORK. Even *restlessness* 6 became more a matter of emotionality, less of "badness." Interestingly, at this same time the trait *dominance* 8 shifted away from "badness" toward good school work. But the second spring ratings showed that Miss A returned to her former position to a real extent. *Dominance* 8 again approached badness with a little PLACIDITY thrown in. The concept of GENERAL EMOTIONALITY disintegrated, to some extent, leaving by itself *lack of emotional control* 9. The trait *emotionality* 5 now became a mixture of all three characteristics. But *restlessness* 6 and certainly *impulsiveness* 11 alined themselves neatly with *misbehavior* 1 to define the pole of BAD BEHAVIOR.

These shifts in time in the configuration presented by one teacher's ratings could represent a number of possibilities.

(1) There could have been a shift in the nature of child behavior itself, from the second to the third grade, so that the nature of trait organization followed a developmental pattern. The tendency for the two spring-semester ratings to agree in *some* respects better than did either of these with the fall rating suggests that there may have been a seasonal variation of some sort; this could have been a variation in the "true behavior" of the children. Such variations would need to be in the basic structure of personality itself; they would have to alter the nature of the intercorrelation between traits.

(2) There might have been a shift in the school environment (including the teacher) such that the trait-correlation shifted. For example, the summer vacation may have brought out a dominance in the child quite compatible with the GOOD SCHOOL WORK syndrome, while the course of the school year may have been such as to push the dominant child over toward the "bad" children, through frustration, let us say.

(3) The teacher's own psychology may have shifted, so that she made her ratings differently. In the fall semester, for example, full of the challenge of a new school year, she may have been more tolerant, more willing to consider dominance a virtue. As the year wore on, she may have reverted to some more basic conviction that dominance was undesirable.

These possibilities are only suggestions, and there may be others. The controls we have used are at best sometimes questionable. We feel, however, that this scheme of attack presents a discriminating picture of variations in the pattern of rater-function in the evaluation of personality. It gives at least some added clue to the nature of factors, like the halo effect, which affect the reliability of measurement. Finally, it suggests a means whereby the functioning psychology of the individual (rater in this case) may be analyzed for its basic framework.

A STUDY OF CHILDREN'S HEARING AS IT RELATES TO READING*

HELEN KENNEDY
Pasadena City Schools

HEARING AS A RELATED FACTOR IN READING DISABILITY

In recent years the question of reading and reading difficulties has occupied an important place in educational and psychological literature. Probably no question has been raised with greater frequency than that concerning the causes of reading deficiencies found with apparently unhandicapped school children. A number of possible causes have been suggested and studied more or less thoroughly. Such suggested causes have included practically every phase of human adjustment —emotional, intellectual, physical, environmental, and pedagogical.

In 1932 Marion Monroe published one of the most significant of the books dealing objectively and specifically with the problems of reading difficulties.[1] This book has been followed by other books which also deal with the causative factors and remedial procedures in reading disability. Many of the recent general books on reading contain at least one chapter devoted to the problems of diagnosis and remediation of reading difficulties.

Certainly no author at the present time claims that any one factor is the sole cause of reading disability but rather the concept of multiple causation is being given greater emphasis. Among the factors which are proposed as possible contributors to reading difficulty hearing deficiency is referred to frequently. Monroe mentioned it but cited no evidence beyond stating that on a whispered voice and coin click tests 2 per cent of the reading defect cases in her study were rated as defective in hearing.[2]

Gates,[3, 4] Betts,[5] Monroe and Backus,[6] Witty and Kopel,[7] and Strang,[8, 9] also mention hearing as a possible contributing factor to reading difficulty and all cite as their only supporting evidence a study by Bond.[10] Cole,[11] and Stanger and Donahue,[12] Durrell,[13] Kirk,[14] and Dolch[15] mention hearing but give no direct reference to any study showing a relationship between hearing and reading, although they do make reference to two of the authors mentioned above—namely Betts and Gates. Only two, those by McCallister[16] and Dearborn,[17] of fifteen of the most recent books on reading or reading difficulties fail to mention hearing deficiency as a possible related factor.

Since it is well known that deaf children experience great difficulty in learning to read and because so many of the writers in the field of reading have assumed a relationship between hearing and reading to exist but have cited little or no evidence as the basis for their assumption it is the purpose of this study to investigate the relationship between hearing and reading achievement of children in a regular school situation.

The relationship between reading and hearing is more or less directly dependent upon the definition which is accepted for reading. The definition is broadening rapidly but there are still wide differences of opinion as to what is involved in the reading process. The definitions vary from the consideration of

* This article is a brief of the author's doctoral dissertation at the University of Chicago.

[1] Marion Monroe, *Children Who Cannot Read*. Chicago: University of Chicago Press, 1932.

[2] *Ibid.*, p. 95.

[3] Arthur I. Gates, *Improvement of Reading*, p. 376. New York: Macmillan Company, 1935 (revised).

[4] *The Teaching of Reading: A Second Report*, pp. 399–400. National Society for the Study of Education, Thirty-sixth Yearbook, Part I. Edited by Guy M. Whipple. Bloomington, Illinois: Public School Publishing Company, 1937.

[5] Emmett Albert Betts, *Prevention and Correction of Reading Difficulties*, pp. 191–99. Evanston, Illinois: Row, Peterson and Company, 1936.

[6] Marion Monroe and Bertie Backus, *Remedial Reading*, pp. 13, 19, and 159. Cambridge, Massachusetts: The Riverside Press, 1937.

[7] Paul Witty and David Kopel, *Reading and the Educative Process*, pp. 241–43. Boston: Ginn and Company, 1939.

[8] Ruth Strang and Florence Rose, *Problems in the Improvement of Reading in High School and College*, p. 209. Lancaster, Pennsylvania: Science Press Printing Company, 1938.

[9] *Reading in General Education*, pp. 310, 312. Compiled and Edited by W. S. Gray. Chicago: University of Chicago Press, 1940.

[10] Guy Bond, *The Auditory and Speech Characteristics of Poor Readers*, Teachers College Contribution to Education. New York: Teachers College, Columbia University, 1935.

[11] Luella Cole, *The Improvement of Reading*, pp. 289–90. New York: Farrar and Rinehart, Incorporated, 1938.

[12] Margaret Stanger and Ellen K. Donohue, *Prediction and Prevention of Reading Difficulties*, p. 22. New York: Oxford University Press, 1937.

[13] Donald Durrell, *Improvement of Basic Reading Habits*, p. 282. New York: World Book Company, 1940.

[14] Samuel A. Kirk, *Teaching Reading to Slow Learning Children*, p. 152. Boston: Houghton Mifflin Company, 1940.

[15] Edward W. Dolch, *A Manual for Remedial Reading*, pp. 11–14. Champaign, Illinois: The Gerrard Press, 1939.

[16] James M. McCallister. *Remedial and Corrective Instruction in Reading*. New York: D. Appleton-Century, 1936.

[17] *Recent Trends in Reading*, pp. 103–10. Compiled and edited by W. S. Gray. Supplementary Educational Monograph, No. 49. Chicago: University of Chicago Press, 1939.

reading as the recognition of the printed word symbols to the recognition of the total process of making discriminating reactions to any stimuli. If the latter definition is accepted then hearing is one form of reading. If, on the other hand, a narrower definition of reading is accepted hearing and reading become two different yet related areas in symbol interpretation. For the purposes of this paper only that phase of the process involving the interpretation of printed word symbols will be considered. It becomes logical, then, that reading, thus defined, is the third step in the normal sequence of symbolization. The first two steps involve the recognition and later the reproduction of the verbal-auditory symbols. In instances of blurred or indistinct auditory impressions it may be difficult to make the associations with the visual symbol. Reading may be further handicapped since an auditory difficulty would limit certain types of direct experiencing as well as vicarious experiencing through normal verbal patterns of communication. On the other hand there may be instances in which this handicap, like myopia, may, in limiting one's normal patterns of learning, cause him to turn more and more to printed symbols as a vicarious form of experience and companionship.

The primary purpose of this investigation is to study the relationship between hearing and reading ability in a regular school situation where hearing losses may exist but in which the special methods of teaching reading employed in the schools for the hard-of-hearing will not have been used. Differences in the type of equipment used, the experimental technique employed, the methods of scoring, and the variations in the age ranges of the children studied are conceived, in part at least, to be responsible for the disagreement in the conclusions arrived at by the earlier investigators. From a consideration of the literature in the fields of psychology, medicine, physics, and sociology as well as education, relating to the problems of hearing it becomes evident that relatively little is known about children's hearing. In the earlier educational studies such a factor as socio-economic status was disregarded although workers in other areas found a definite relationship between hearing ability and the socio-economic conditions of the homes from which the children came.

In order to investigate the relationship between reading and hearing, then, it becomes necessary first to make an investigation of children's hearing. For the purposes of this study 433 individuals between the ages of six and twenty-three were given an audiometric examination with a Western Electric continuous frequency audiometer. All of these persons were from the University of Chicago or the Laboratory Training School of the University. The population cannot be considered as typical but rather were selected for the relative socio-economic homogeneity thus assured and the availability of additional data.

Two aspects of children's hearing were investigated—acuity and discrimination. Using the data obtained from this investigation the study was carried further to include the relationship between these factors and silent and oral reading ability. All the children in the first, second, and third grades were given both silent and oral reading tests. The children in the fifth, seventh, and tenth grades were tested only in silent reading. Intelligence test data were also available for all students for whom there were reading data.

Review of the Literature

After an extensive review of the literature relating to hearing several differences were identified between the studies of the relationship between hearing and reading which might account for the variances in the conclusions arrived at in the five investigations of this relationship.

Bond[18] found a definite relationship between hearing and reading. Aretz[19] found a slight relationship while Brock[20] and Madden[21] found no relationship. Conway[22] reported some relationship where the hearing loss[23] was greater than twenty decibels. The most important variations in these studies occurred in respect to (1) the consideration

[18] Bond, *op. cit.*

[19] John L. Waldman, Francis A. Wade, and Carl W. Aretz, *Hearing of the School Child.* Philadelphia: Temple University, 1930.

[20] Elma E. Brock, *The Relationship Between Auditory Acuity and Achievement in Silent and Oral Reading.* Unpublished Master's Thesis, University of Chicago, 1939.

[21] Richard Madden, *School Status of the Hard-of-Hearing Child,* Teachers College Contribution to Education. New York: Teachers College, Columbia University, 1931.

[22] Clifford Conway, *The Hearing Abilities of Children in Toronto Public Schools.* Bulletin No. 9. Ontario: Department of Educational Research, College of Education, 1937.

[23] Hearing loss is a term which through common usage has come to refer to any variation from the "normal." Thus there may be a minus hearing loss which would mean hearing acuity better than "normal."

given to intelligence, (2) procedures in the selection of the subjects, (3) the age differences of the subjects, (4) the type of equipment used, and (5) the methods of scoring. Further differences were noted in the studies with regard to sex.

From findings in the fields of psychology, medicine, physics, sociology, and music at least partial answers may be given in some of these areas of differences. Further, other factors not considered are found to be important. Socio-economic conditions are found to play a very important part in the prevalence of hearing loss yet this condition received little or no consideration in the five studies mentioned above. The possibility that hearing may vary with age in children as well as with adults has little direct supporting evidence but a wealth of suggestive data are available.

Another factor of importance in an understanding of the problem relates to the type of equipment used. The two kinds of testing equipment used most generally are not comparable since the group testing equipment (4-type) gives no indication of a high frequency loss. These high frequency losses are the variety found by Ciocco[24] and Conway[25] to be particularly prevalent among boys and account for a high percentage of all hearing loss. From the consideration of data in the field of physics it is clear that the study by Gage,[26] accepted by Bond[27] and Brock,[28] is in error and that binaural hearing cannot be considered to be better than the hearing in the better ear. This indicates that the score for the better ear only is not comparable to some form of score which considers both ears. From the other data, then, it appears that the differences in the methods of scoring and the kinds of equipment used by the experimenter in the field of education and its relationship of hearing may account for the lack of conclusiveness in the findings. It seems possible, also, that the failure on the part of these experimenters to take into consideration certain factors such as age and socio-economic status, at least in part, may have distorted the evidence.

[24] Antonio Ciocco, "Audiometric Studies in School Children, V," *Annals of Otology, Rhinology and Laryngology*, XLVII, pp. 926–37.

[25] Conway, *op. cit.*

[26] F. H. Gage, "A Note on Binaural Threshold," *British Journal of Psychology*, XXXIII, (1932), 148–51.

[27] Bond, *op. cit.*

[28] Brock, *op. cit.*

The evidence regarding auditory discrimination is even more meager than that of auditory acuity. Most of the available evidence comes from the field of music. Much of the available data concerning the possibility of maturation of auditory discrimination is controversial although the more recent findings tend to favor such an hypothesis. The usual educational technique used in measuring auditory discrimination consists of dictating pairs of words and the subject is asked to indicate whether they are the same or different. These tests lack evidence of reliability or validity as well as being very short and containing a variety of changes such as the addition or omission of sounds, reversals of the sound patterns as well as certain substitutions of sounds.

From these findings it becomes evident that a more careful and conclusive study must be made of the hearing of a group of children of various age levels before its relationship to certain academic pursuits can be established. This, then, becomes the first consideration of this study—the study of the hearing of a large group of children of various age levels.

Auditory Acuity

As has been stated all the subjects were drawn from the Laboratory Training Schools of the University of Chicago or from the University itself. This assured a homogeneity in the socio-economic status of the homes from which the children came. It also assured the availability of additional data regarding the subjects since a records office and testing personnel are maintained by the schools. The smallest number of any group of subjects on the basis of either age or grade consideration is forty-five. From this number the groups vary up to the largest group which contains seventy-three. As in most groups considerable variation was found in intelligence. The subjects in the lowest five grades had been given Stanford Binet Intelligence tests. The medians for the different grades varied from an intelligence quotient of 121 to 128. The range for individuals was from 97 to 179, obviously a very superior population. At the seventh grade level the Kuhlman Anderson test was used with a variation from 92 to 158, with a median of 119. The American Council on Education Psychological Examination was used with the

tenth grade. The total scores, in percentile rankings, were from 3 to 87, with a median of 36.

The conditions under which the audiometric tests were conducted were not ideal but because of comparative isolation the room in which the tests were given was exceedingly quiet. All the subjects were tested in the same room by the investigator. In finding the minimum audible threshold for each individual the approach was made from below the level of audibility. The subject was allowed to hear the tone at approximately twenty decibels (db.) or higher if necessary. Using the cut-off switch the dial was turned to the minimum intensity or to at least twenty db. below his expected minimum level. He was then asked to indicate the point at which the sound appeared when thus approached from below the threshold. The first test was made at 512 double vibrations (d.v.). The procedure was repeated at each octave level to 8192 d.v. and again at 9747 d.v. A retest was made at 512 d.v. and the two lower octaves were then tested. The scale was then run as near the minimum threshold as possible in order to check for tonal gaps. This could be done quite accurately by the manipulation of both the volume and frequency dials and the cut-off switch.

The typical audiogram was not used but one was made by the writer in which the frequency levels on the abscissa were represented on a logarithmic scale which corresponded to the scale found on the frequency dial. This made possible a more accurate recording of losses.

Several types of studies of the auditory acuity data were made. Fifty-five of the first and second grade pupils were retested with the first ear tested and a reliability study made at each of the eight frequency levels. These data are presented in Table I.

A second coefficient of correlation was obtained at 128 d.v. using only fifty-three of the subjects. Two cases were discarded because their second test deviated greatly from the first at the low frequencies. With this change the coefficient of correlation was raised to $.70 \pm .047$. The writer believes it probable that the low correlations at the low frequencies do not indicate so much a lack of reliability in the tests as an actual change in hearing. In some instances as much as three months time intervened between the test and retest which in a number of instances was due to illness. Since this testing program was conducted during the winter months and colds were prevalent it seems probable that there may have been temporary changes in hearing at the low frequencies. The fact, too, that the high frequencies had much higher correlations would seem to lend support to this belief.

Several other correlation studies were made. At 128 d.v., with 428 cases, a coefficient of correlation of the right with the left ears was found to be $.585 \pm .022$. The results of the first ear tested were correlated with the results for the second ear and found to be $.573 \pm .022$. A third correlation made with all these cases was that of the hearing acuity in the right ear with age. The coefficient of correlation was $.388 \pm .028$.

Another type of correlation study was made with 207 boys. The auditory threshold for the right ear at each frequency level was correlated with the threshold at each of the other frequencies. From these data it is seen that the relationship decreases as the differences in the frequencies increase. The highest correlation, $.795 \pm .017$, is that between 128 d.v. and 256 d.v. The lowest. $.217 \pm .044$, is found between 256 d.v. and 9747 d.v. Even though all the correlations are statistically significant it is evident that one cannot predict the auditory acuity for any frequency level from any other level. Thus it is clear that any attempt to obtain "a" hearing score, even for one ear, would be misleading and might disguise real points of significance in a study.

One of the primary hypotheses of this study concerns the possibility of a growth or

TABLE I

COEFFICIENTS OF CORRELATION AND PROBABLE ERRORS OF COEFFICIENTS BETWEEN FIRST AND SECOND MEASURES OF AUDITORY ACUITY

d.v.	128	256	512	1024	2048	4096	8192	9747
r	.561	.680	.597	.786	.807	.905	.854	.719
PE_r	.062	.049	.058	.036	.033	.017	.024	.044

maturation factor in auditory acuity. In view of this hypothesis the means and standard deviations of the minimum auditory threshold for each of the seven age groups are presented in Tables II to V, inclusive. Data are presented in terms of each of the eight major frequencies indicated on the audiometer. The presentations are made in separate tables for the boys and girls and for the right and left ears. Because the groups were necessarily small when separated to this extent the medians were also calculated. The interpretation of these data will be made on the basis of the means, however, since the differences between the means and medians were slight. Any deviations in the pattern of the curves for the means and medians will be noted.

These tables may be considered in either of two ways. If they are read horizontally it becomes evident that at no age level do the means approximate "normal"—the zero line. The general configuration of the curves, however, are more or less uniform—the two lowest frequencies are markedly lower than the others, the curve rises rather abruptly at 512 d.v., and tends to straighten above that level. There is a tendency for the curve to drop again at the high frequencies with the six-year-old and the adult groups. It would seem from these data that additional adjustments need to be made in the instruments but these findings do not, of course, affect the findings of the study.

If these tables are read vertically age differences become apparent. It is not enough, however, to be able to point out that certain differences occur but rather it is necessary to know whether these differences are significant. The significance of all the differences between age, sex, and laterality groups have been studied and may be summarized as follows:

1. There is a highly significant difference between the means of the six- and fifteen-year-old students in practically every instance. The only exceptions to this are at 4096 d.v. with the boys, with both the right and left ears, and at 2048 d.v. with the boys, using the left ear.

2. At almost every frequency level there is a significant difference between the six- and eight-year-old groups. This holds for five of the eight frequency levels for the right ears of the boys and six for the girls. It is significant for all frequencies with the boys' left ear and all but one frequency with the girls' left ear.

3. The differences between the eight- and fifteen-year-old groups are significant for the boys in less than half the instances; but with the exception of four cases, two right and two left, all these differences are significant for the girls.

4. It is interesting to note, although the differences are not statistically significant, that the mean hearing for the ten-year-old boys is in every instance less acute than for the eight-year-old boys. This is not true at any frequency level for the girls. On the other hand, they show a mean loss at the twelve-year-old level as compared with the ten-year-old group. Some of these losses are significant. It is true, however, that the differences between these groups of boys practically disappear when medians are considered rather than means. Using medians, the lower acuity of the twelve-year-old girls is equally as marked as with the means above 1024 d.v. In some instances a similar drop occurs in the medians for the boys.

5. An even more striking consistency is a decrease of the means and medians between the fifteen-year-old and adult groups. In the boys' right ear and the girls' left ear there is a slight increase in the means at 4096 d.v. and with the right ears of the boys at 8192 d.v.

6. Obviously the frequency level 4096 d.v. needs special study since the changes here vary from the general pattern. At this level the gains are insignificant when increases are significant for all other frequencies. Similarly when there are decreases at other frequencies there is a slight gain at 4096 d.v. The writer has found in the literature and from conferences with otologists that this is a level of particular interest to the medical profession.

7. On the whole there seems to be more growth, and more significant growth, among the girls than among the boys. This is particularly true in the lower and middle frequency ranges, but is not true at the highest frequency where the boys have a more significant growth.

8. While few of the differences prove to be statistically significant it is interesting to note that the boys, on the whole, are better in the younger groups and that there seems to be a definite superiority of the girls in the older groups. There is also a greater

TABLE II

MEANS AND STANDARD DEVIATIONS (σ) OF THE AUDITORY ACUITY FOR EACH OF EIGHT FREQUENCIES FOR THE BOYS' RIGHT EAR

	128		256		512		1024		2048		4096		8192		9747	
Age	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ
6	5.95	5.49	5.32	4.29	— .59	4.11	—2.91	4.66	—5.50	4.67	— 6.45	3.98	.91	15.52	1.51	13.68
7	2.78	5.18	3.55	4.72	—1.89	4.53	—4.33	3.69	—6.78	3.24	—7.78	4.66	—6.45	7.17	—6.45	7.56
8	2.92	7.09	2.13	6.03	—3.19	4.89	—5.62	4.93	—7.69	4.35	—7.46	5.25	—7.96	4.47	—5.50	6.02
10	2.95	10.40	3.50	9.48	—2.77	11.40	—4.82	13.14	—7.00	12.91	—5.09	13.65	—3.18	15.13	—1.95	17.75
12	— .77	4.41	1.77	6.12	—3.88	3.48	—5.27	2.91	—8.39	4.28	—7.35	9.87	—5.76	18.95	—6.65	10.27
15	— .69	4.86	— .48	4.92	—6.69	3.18	—9.79	3.05	—10.21	4.41	—6.48	15.93	—9.28	10.47	—10.62	5.60
Adult	.50	3.81	2.41	5.52	—4.41	3.97	—8.09	3.79	—8.64	6.29	—8.77	6.37	—9.45	7.33	—6.87	9.54

TABLE III

MEANS AND STANDARD DEVIATIONS (σ) OF THE AUDITORY ACUITY FOR EACH OF EIGHT FREQUENCIES FOR THE BOYS' LEFT EAR

	128		256		512		1024		2048		4096		8192		9747	
Age	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ
6	8.55	5.33	8.55	4.67	— .32	3.76	—3.05	4.65	—4.41	2.76	—4.27	5.11	.50	6.94	2.27	8.91
7	3.78	5.72	3.78	6.01	—1.67	4.28	—5.33	4.03	—7.56	4.90	—7.67	5.18	—5.56	7.76	—3.11	6.36
8	3.73	5.67	2.81	5.62	—3.54	4.29	—6.08	4.39	—8.27	4.87	—7.81	6.68	—5.85	7.36	—5.15	8.89
10	5.23	14.12	6.36	15.39	1.32	19.40	—2.64	14.41	—5.50	14.22	—3.05	19.98	— .73	16.85	—1.95	16.18
12	.96	4.88	1.08	4.92	—2.39	3.24	—5.96	4.47	—6.55	5.73	—9.08	5.76	—7.92	11.11	—2.96	10.22
15	.66	5.08	1.38	5.86	—5.34	3.75	—7.83	4.23	—7.21	9.96	—5.66	13.54	—11.24	9.24	—9.28	8.52
Adult	4.59	14.53	5.96	15.50	—1.15	11.72	—5.37	10.02	—4.49	12.72	—4.41	15.58	—5.64	12.42	—5.37	13.09

TABLE IV

MEANS AND STANDARD DEVIATIONS (σ) OF THE AUDITORY ACUITY FOR EACH OF EIGHT FREQUENCIES FOR THE GIRLS' RIGHT EAR

	128		256		512		1024		2048		4096		8192		9747	
Age	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ
6	8.78	5.70	9.96	5.78	.83	4.59	—2.70	3.52	—5.83	4.17	—5.70	5.22	—2.17	8.97	— .74	9.90
7	5.21	6.31	6.18	6.96	— .69	4.83	—4.32	4.20	—5.71	3.71	—5.71	6.58	—3.68	7.45	—3.57	10.06
8	1.55	5.95	1.85	5.57	—2.95	3.71	—7.15	4.69	—7.60	4.71	—7.90	4.85	—7.30	5.76	—6.21	6.23
10	— .14	4.46	.50	4.43	—4.64	4.64	—7.54	5.27	—9.79	4.41	—11.39	5.86	—6.79	12.09	—6.36	10.91
12	— .40	3.63	— .10	3.69	—4.75	2.38	—5.95	2.89	—6.70	2.83	—8.65	2.96	—8.65	6.41	—6.55	6.81
15	— .72	3.69	—1.19	3.28	—6.16	3.29	—8.41	4.74	—10.66	3.95	—11.69	3.64	—11.69	7.31	—8.78	9.30
Adult	.56	6.90	1.64	6.48	—4.84	5.40	—7.60	6.18	—9.16	5.72	—8.84	10.62	—6.52	12.10	—4.12	12.49

TABLE V

MEANS AND STANDARD DEVIATIONS (σ) OF THE AUDITORY ACUITY FOR EACH OF EIGHT FREQUENCIES FOR THE GIRLS' LEFT EAR

	128		256		512		1024		2048		4096		8192		9747	
Age	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ
6	10.61	6.90	10.61	6.69	1.76	5.03	—1.78	4.35	—4.91	5.48	—4.39	6.40	— .22	6.99	1.61	9.92
7	5.96	5.61	6.61	6.13	—1.75	3.55	—4.43	4.92	—6.57	4.22	—7.11	5.10	—2.50	6.29	—3.04	8.51
8	3.20	5.89	3.20	8.13	—2.65	5.49	—5.95	4.78	—8.65	4.79	—6.55	6.24	—5.65	7.50	—5.80	8.45
10	1.89	5.82	1.57	6.79	—4.54	4.34	—7.54	4.17	—9.46	5.07	—10.11	4.78	—7.21	8.22	—9.14	8.49
12	1.10	3.69	1.70	3.90	—4.00	3.42	—5.65	3.23	—6.10	4.47	—7.45	4.27	—4.00	5.70	—2.50	8.25
15	—1.00	3.90	—1.09	3.48	—6.34	2.97	—9.63	3.96	—10.47	4.02	—10.19	7.36	—12.53	7.25	—11.59	7.50
Adult	.08	5.43	.80	6.06	—5.80	3.80	—8.92	2.80	—10.16	3.88	—10.84	3.84	—7.60	9.90	—6.88	7.86

superiority in favor of the girls with the left ear than with the right.

9. The medians do not show the same variations between the sexes at the ten-year-old level that is evident in the means. This very lack of difference in the two measures of central tendency, however, is evidence to substantiate Ciocco's findings that there are much greater losses among the boys, particularly at the ten-year-old level, than among the girls.

10. There is a general superiority of the right ear as compared with the left with all the groups of boys and for the younger girls. The fifteen-year-old girls and the women have a definite left ear superiority. The only hypothesis which the writer has to offer for this apparent shift from greater right to left ear acuity is the possibility that hearing is functional and that this change may in some way relate to the earlier and more frequent use of the telephone by women than by men. Possible support for this hypothesis may be found in the reports of the hearing of telephone operators.[20]

Auditory Discrimination

Although hearing acuity is basic to the understanding of speech it is not the only factor which operates in this regard. It is not only necessary to be able to hear sound but one must be able to distinguish between the various sounds. Because the second factor is important such a study as this could not be completed without some consideration of auditory discrimination. Each of the speech sounds has a basic fundamental frequency which differentiates it from every other speech sound. Between certain speech sounds there are very slight differences while there are great differences between other sounds. For example, "f" and "v" are both high frequency sounds and consequently are much harder to distinguish between than "f" and "n". In other words it is harder to hear the differences between "fast" and "vast" than between "fear" and "near".

Three techniques were used in an attempt to obtain data concerning the individual's ability to discriminate between given sounds. The Seashore Pitch Discrimination test was used on a 4–C audiometer thus assuring that each child received the same stimulus. It was not possible to administer this test to all of the children for whom other test data were available and no attempt was made to use this test below the fifth grade. The findings, therefore, are based on data obtained for the ten-year-old boys and the boys and girls in the twelve- and fifteen-year-old groups. When comparisons were made between the ten- and twelve-year-old and between the twelve- and fifteen-year-old boys the differences were found to be significantly higher in the older group. The differences between the twelve- and fifteen-year-old girls favor the older group but are not statistically significant.

A second technique for testing pitch discrimination was employed with the 6–B audiometer but was not found to be satisfactory. The intensity dial was set forty decibels above the minimum audible threshold for the individual at 512 d.v. and each of the octave levels above. By using the cut-off switch the writer attempted to determine the minimum change in frequency which the subject could identify as a change. The general pattern of differences with both the boys and the girls was an increase in this ability with age, but these were not statistically significant.

The third type of test which was used in an attempt to find differences in the children's ability to discriminate between given sounds was patterned after the articulation tests used for testing the efficiency of telephone circuits and the acoustical properties of rooms. None of these tests was given to the subjects of the general experiment but will be described briefly since these pretests seem to offer greater possibilities than either of the other discrimination techniques employed. These tests are based on a sound pattern composed of an initial and terminal consonant and an internal vowel. When first tried the list was composed of words which had as low a count on the Gates and Thorndike Word Lists as possible, a high count on the Thorndike List, and a series of these sound elements which did not make words. Because the familiar words were recognized without difficulty and the others were frequently confused a second test composed entirely of nonsense syllables was tried. The elements composing each of these units was purely chance having been drawn from piles of cards containing the sounds. This test was administered to a group of twenty-six eighth and ninth grade students and all the conson-

[20] Dana W. Drury, "Tests of Hearing of 500 Average Ears by the Audiometer, Two-A." *Archives of Otolaryngology*, I (May, 1925), pp. 524–32.

ant errors were recorded. It was found that 30.98 per cent of all the consonant sounds thus dictated were missed and of these two-thirds of the errors were of such a nature as to form a common word. These findings seem to the writer to definitely point to the conclusion that sound combinations are perceived as wholes. While much more work is needed in this area this technique seems to give promise of more valid and reliable results than any of the other techniques now employed.

Reading and Its Relationship to Hearing

To the educator the significance of the findings reported above lies primarily in the relationship which may be found to exist between hearing and academic accomplishment, one aspect of which is reading. In order to study this relationship silent reading test data were obtained for the first, second, third, fifth, seventh and tenth grade children and oral reading data for the first three grades. The Standardized Oral Reading Paragraphs were used for the oral reading data. The reading section of the Metropolitan Achievement Test was used in the five lower grades studied and the Iowa Silent Reading Test was used in the tenth grade. These tests were administered by personnel of the Laboratory Schools and the results made available to the writer.

At the outset of the study of the relationship of these factors—hearing and reading—it must be recognized that this study can be no more than suggestive because of certain limitations which exist. First, the cross-sectional nature of the study does not permit the identification of the subjects who may have had a hearing loss for some time in contrast with those whose loss is recent and may be temporary. Furthermore it is not possible to determine the severity of the loss at the onset of the difficulty. Second, the superiority of the subjects academically, intellectually, and probably auditorially is a limitation imposed on this part of the study, but which operated in favor of the study of possible growth factors in hearing.

The superiority of the groups in reading in relation to national norms is clearly evident from the data presented in Table VI.

Since the reading tests in all but the tenth grade were administered in May the expectancy for each of the grades would be about nine months above the grade level. For example the second grade should be about 2.9 whereas actually they are nearly one year accelerated in silent reading and one and one half years in oral reading. The remainder of the table is read in the same manner.

In attempting to treat the hearing data in relation to reading the use of sixteen different hearing scores becomes confusing and impractical. Different workers in the field have used different types of composite scores, none of which seems satisfactory. It seems that true binaural hearing must represent some condition less acute than the hearing in the better ear yet superior to the hearing in the poorer ear. What this relationship may be apparently is unknown. The writer has, therefore, added the scores for the eight frequencies in each ear and is using that as a composite score. This is obviously a very crude method but does represent the entire range and both ears and does show marked differences among individuals. A zero score approximates the so-called "normal." A minus score would mean better than "normal," while a plus score denotes hearing less acute than "normal." The composite hearing scores for the various grade levels are given in Table VII.

These data vary slightly from those presented in the earlier tables. These are presented by grades rather than age levels because it is in this connection that the rela-

Table VI

Reading Test Data

Grade	Silent Reading No.	Q_1	Median	Q_3	Oral Reading No.	Q_1	Median	Q_3
1	37	2.64	2.14	1.86	41	3.09	2.39	2.06
2	52	4.50	3.79	3.13	54	5.17	4.42	4.49
3	47	6.51	5.53	4.54	52	6.95	5.18	5.49
5	59	8.03	7.56	7.11	----	----	----	----
7	61	9.96	9.50	8.90	----	----	----	----
10	72	76*	59*	41*	----	----	----	----

* Percentile scores.

TABLE VII

COMPOSITE HEARING SCORES

Grade	Number	Q_3	Median	Q_3	Range
1	46	— 41.0	+ 7.5	+ 47.5	— 97 to +214
2	57	— 79.0	— 41.0	— 10.0	—175 to +109
3	52	—109.0	— 73.0	— 32.0	—198 to +229
5	62	—125.0	— 76.0	— 21.0	—198 to +591
7	63	—104.0	— 62.0	— 38.0	—170 to +141
10	72	—162.0	—128.0	— 74.0	—221 to +167

tionship with reading will be studied. From these composite scores the growth factor is clearly evident.

Since the main objective of the study is to examine the relationship between reading and hearing a correlational study between these two factors was made at each grade level. While all the correlations, except for the second grade with oral reading, are positive they are too low to be significant. The highest is for the tenth grade, $+.344 \pm .071$. This is not surprising, however, since no condition other than grade level has been held constant yet many other factors are known to affect reading. It is necessary then to study further the relationship between reading and hearing by other means.

It is evident from the reading data that the average reading ability of the pupils used in this study is considerably above normal; there are within the groups, however, wide differences in reading ability. One method of considering the relationship between hearing and reading is to take the highest and lowest 15 per cent in reading ability at each grade level and study their hearing. A second method of studying this relationship is to select groups on the basis of their hearing ability and study their reading ability. This selection was also made after a careful study of the audiograms and scatter diagrams. Approximately 12½ per cent at each grade level were thus selected. An effort was made to pair the poor hearing students with superior hearing students of the same age and sex. Intelligence was also considered but because of the small number of cases at each grade level this basis was necessarily abandoned.

Although there was a somewhat consistent difference in favor of the means of the better hearing or better reading group at each grade level these differences were not found to be statistically significant. It will be recalled, however, that in the introduction two hypotheses were suggested: (1) that a hearing handicap would so limit a child's contacts with the world that this lack of experience would be reflected in his failure to achieve normally in the interpretation of the printed page, or (2) that because of the limitation in direct experiencing imposed by an auditory handicap, the child might turn to the printed page for a vicarious form of experiencing and thus, through reading, become a better reader than his unhandicapped classmate. Thus no study of the relationship between reading and hearing can be complete without a consideration of the normalcy of the distribution.

Scatter diagrams were made showing into which quartile of reading or hearing ability each child belonged in relation to the other members of his grade. These scatter diagrams were made for the reading groups, the hearing groups, and for the total population of the study. In Table VIII are presented the scatter diagrams for the hearing and reading groups. From these data it is obvious that there is a significant difference in the distribution of hearing scores of the better and poorer groups in silent reading but not in oral reading. Similarly there is a significant difference in the distribution of the silent reading scores of the good and poor hearing groups. For the hearing groups it is questionable whether the difference in the distribution of the oral reading scores is really significant. The score is of such a nature that it must be interpreted with caution. It is probable that the hypothesis of a significance in the distribution of these scores should neither be accepted nor rejected without further evidence.

The fact that there is a significant difference in the distribution of the scores although there is not a significant difference in the means is probably due to a cumulative effect in the latter instance since in studying the differences of the means the grade-level groups were considered separately while the findings were combined for the consideration of the distributions since there were only six pupils at several of the individual grade

TABLE VIII

DISTRIBUTION OF THE UPPER AND LOWER READING AND HEARING GROUPS INTO APPROPRIATE QUARTILES IN AUDITORY ACUITY AND READING

Hearing Quartiles	Reading Groups: Silent Reading 15 per cent Upper	Reading Groups: Silent Reading 15 per cent Lower	Reading Groups: Oral Reading 25 per cent Upper	Reading Groups: Oral Reading 25 per cent Lower	Reading Quartiles	Hearing Groups: Silent Reading Upper	Hearing Groups: Silent Reading Lower	Hearing Groups: Oral Reading Upper	Hearing Groups: Oral Reading Lower
1	16	10	10	10	1	16	8	7	3
2	15	9(10)	13	9(10)	2	10	9	4	3
3	11	13(14)	8	7(8)	3	6	11	3	7
4	11	19	8	11	4	8	12	4	5
Total	53	51	39	37		40	40	18	18
Chi-square	10.74		2.74			11.71		8.15	

levels, an inadequate number for the consideration of the normalcy of the distribution.

The distribution of both the upper and lower groups are skewed but in the opposite directions. This same tendency exists when all in the upper and lower quartiles are considered rather than the smaller selected groups. When the quartiles were used there were eighty-two and eighty-four cases respectively for the reading and hearing groups. It is interesting to note that a child of good hearing ability apparently has a two to one chance of being a superior rather than a poor reader.

When the data were further broken down it was found that of the eighteen children with the lowest hearing ability in the first three grades only three made oral reading scores in the upper quartile and all of these were in the third grade. No first or second grade child with relatively poor hearing was one of the upper twenty-five per cent in oral reading.

While these data do not prove conclusively that there is a relationship between reading and hearing they do prohibit the discarding of such a hypothesis without further investigation. It may be that the groups used in this study were too highly selective to permit the true relationship to fully manifest itself. It is also recognized that hearing can be only one of the many factors which may contribute to success or failure in reading. Since relatively so little is known of children's hearing it was deemed advisable to study further the records of the poorer hearing group to determine whether there might be evidence to indicate why some of these children read very well and others experience considerable difficulty in reading.

Age and type of hearing loss were two major considerations in this additional analysis. The relationship between reading and hearing loss seems to be related to grade level. The poorer hearers in the fifth and lower grades tend toward superior reading, while in the seventh and tenth grades the chances are about four to one that a poor hearer will fall below the median of the class in reading.

The type of hearing loss, however, seems to be the real determining factor in the relationship of hearing and reading. The seven with only a low frequency loss were better readers while the eight with high frequency losses tend to be poor readers. Of the low frequency losses three were in the first quartile, two in the second, and one each in the third and fourth. The high frequency losses were distributed very differently—two in the upper quartile, none in the second, two in the third, and four in the lowest. Low frequency losses tend to be most common in the lower grades while high frequency losses occur in the upper grades. General losses are more or less evenly divided among the upper and lower grades. There is a slightly greater tendency, two to three, for those afflicted with general losses to have reading scores below the group median. High frequency losses seem to be more severe than either the general or the low frequency hearing losses.

Of the other factors considered—intelligence, school absence, vision, and weight—weight alone seems to deviate and that only in cases with a high frequency loss.

In regard to auditory discrimination the silent reading groups differ significantly as this factor was tested with the Seashore records on the 4–C audiometer. These differ-

ences were significant both as judged by the differences in the means and in the distributions of the scores. The discrimination scores for the test made with the 6–B audiometer did not show significant differences between the good and poor readers although they generally favored the good reading group.

Summary and Conclusions

Before interpreting the findings of the study certain limitations which affected them should be discussed. Two factors, the superiority of the subjects and the cross-sectional nature of the study, undoubtedly limited the value of the findings. The group used for this study were superior in many ways. In the first place the intelligence of the subjects was much above normal. The lowest intelligence quotient for any pupil in the lowest five grades studied was ninety-two. From this the range extended to a maximum of 179. The lowest median intelligence quotient for any grade level was 119.

The high socio-economic background of the group offered both advantages and disadvantages in this study. The homogeneity of the background was important in the study of the possible differences in hearing ability which it was thought might relate to age. It also assured a minimum of uncorrected health conditions which affect hearing, such as diseased tonsils and adenoids. While these factors contributed advantageously to the study of development in hearing they offered the disadvantage, from the experimental viewpoint, of there being so few deviating in hearing. While it is known that those extremely handicapped auditorially have great difficulty in learning to read the point at which a hearing loss becomes significant in this respect has not been identified. The auditory superiority of the present group limited the possibilities of making any such identification. Although a group of forty was selected as having poor hearing it is probable that in a more typical population the majority of these children would not have been considered as low.

The group was also found to be superior in reading. It will be recalled that the subjects selected as the lowest fifteen per cent in reading, and hence those who composed the poor reading group, were in many instances normal readers in terms of national norms for the test, but were low in comparison with their own grade level in this particular school. The first and tenth grade groups alone might be considered as containing pupils who were retarded in reading.

A cross-sectional study always offers limitations when one desires to study possible growth factors. Certainly support for the hypothesis that there is maturation or growth of hearing was found. Before the nature of such change can be studied in detail, however, a longitudinal study must be made. The cross-sectional technique is particularly handicapping in a study of the relationship of hearing and reading, since the significance of the relationship may lie not in the fact of a hearing loss but in the period of onset of the loss, its severity at onset, and the duration of the loss.

A third limitation of the study is found in the inadequacy of the number of subjects. While the numbers were adequate in certain respects, they were not sufficient to permit proper matching of pairs for the more detailed study of the possible relationship between hearing and reading.

While these limitations were very real, certain findings may be pointed to as significant. Most important of the conclusions arrived at is that relating to the change in hearing acuity. There are significant differences in the mean auditory acuity of children between the ages of six and fifteen years. The differences are most evident between the six- and seven-year-old levels, the seven- and eight-year-old levels, and between the twelve- and fifteen-year-old levels. No significant differences were found between the eight-, ten-, and twelve-year-old groups.

Two other factors concerning auditory acuity are interesting and may have implications for education. These conditions are consistent although they are not statistically significant. First is the superiority of acuity of the boys at the lower age levels, while the girls exceed at the fifteen-year-old level. Second, there is a right ear superiority with the boys in general and with the girls below the fifteen-year-old level. At the fifteen-year-old level and among the women studied there is a left ear superiority. Might this be related to the use of the telephone?

The discrimination test data contribute two findings worthy of particular notice. First, tests of pitch discrimination seem to substantiate the assumption that there is a

change concurrent with age. Second, from the articulation test data, developed in an attempt to test discrimination with speech sounds, it becomes evident that sound combinations are interpreted as wholes and not as individual elements.

Because of the many limitations of the investigation it has not been possible to reach clear-cut conclusions regarding the relationship between hearing and reading. Correlations of auditory acuity and reading ability while positive are not statistically significant except at the tenth grade level. Even there the correlation is only $+ .344 \pm .071$.

In the study of the selected groups, the means are not found to differ significantly, but the distributions of the scores do vary significantly in silent but not in oral reading. There is a somewhat greater tendency for good hearers to become good readers and for those somewhat handicapped in hearing to read less well. Similarly a good reader is more likely to have good hearing while the poor reader is more likely to have poor hearing.

One of the most interesting findings of the study is in the relationship of the type of hearing loss to reading. When the forty subjects judged to be most severely handicapped auditorially were further divided into groups depending upon the type of their hearing loss it was found that those with a high frequency loss, above 2048 d.v., tended to become very good and more often poor readers. Of the eight thus classified two were in the upper quartile of their grade in reading, two in the third quartile, and the remaining four in the fourth quartile. Only two children in the first three grades were classified as having a high frequency loss and both of them had oral reading scores in the lowest quartile.

The groups selected on the basis of silent reading test scores were found to deviate significantly in auditory discrimination as measured by the Seashore Pitch Discrimination test used on a 4–C audiometer.

With both the findings and the limitations of this study in mind the question of the implications of the investigation must be considered. Certain of the implications which may be pointed to in terms of reading may also be drawn in other areas.

The most striking thing which must be considered is the inadequacy of knowledge relating to the children educators propose to teach. It is well recognized that the major learnings come through two sense avenues, vision and hearing. From the data presented in this study it seems clear that there is a developmental change in hearing, both in acuity and in discrimination. Little or no effort, however, is directed toward the possibility of assisting the child in his recognition of speech sounds and the minor differences between certain of the sound elements. How much of this change is due to education and how much to other factors is one of the many problems still awaiting investigation.

The results of the articulation tests seem to lend support to the concept that one reacts to a whole pattern rather than to the parts. It will be recalled that in the articulation tests, three sounds not making a word were presented at intervals of approximately one second per sound. The results show that two-thirds of all the errors made were in the nature of a substitution of a sound which produced a meaningful sound pattern. In presenting new words it seems logical that the minor sound differences should be made clear to avoid the substitution of a familiar sound pattern for the new one. In teaching reading, then, it seems important that both the visual and the auditory be given attention and that attention be directed to the total configuration, both auditorially and visually, rather than to the individual component elements. Much additional work is needed in this area.

In education the most common test of hearing is the group, 4-type test. It was pointed out that this instrument does not measure hearing in the high frequencies, yet this is the area in which the greatest relationship with reading seems to exist. The clinician interested in determining the existence of possible hearing loss in instances of reading difficulty should find some means of supplementing the test. If the more adequate audiometer is not available at least a set of tuning forks should be used.

Obviously there is no ready answer to the problem of the relationship between reading and hearing. It may be that the data reported in this study are not refined sufficiently to reveal their true value. It may be that hearing is only a part of a larger and more complex system of influences which determine reading ability. If the latter is the case, and the writer believes that it is, the

problem is one of great magnitude and constitutes a challenge which will not yield to superficial attack.

Among the many problems which need further investigation the following seem most urgent.

A longitudinal study of a large group of children is most urgently needed. Ideally such a study should be initiated at the kindergarten or even preschool level. A population large enough to assure the maintenance of perhaps a minimum of fifty at the termination of the study would be desirable. Best results would be obtained only if this population could be followed over a number of years. Only by such a procedure can the hypothesis of this study, that growth does occur in hearing, be substantiated. This would permit determining the relationship of the period of onset, the severity, and the duration of the hearing difficulty with reading. For such a study a typical population which would contain more children with reading and auditory handicaps should be used, not the highly selective group used in this study. In the study of a large typical population additional studies could be made regarding differences relating to socio-economic status. Such a study should only be undertaken where a continuous frequency audiometer is available and arrangements can be made for all testing to be done under fairly consistent conditions of quiet.

For the maximum of value such a longitudinal study would also include the possibility of the study of the relationship existing between hearing and reading and other factors known or thought to affect reading success. Among such factors might be mentioned vision and the deviations in the endocrine functions.

Additional studies which might be made either in conjunction with the larger study or independently include the following:

A study is needed to determine how much of one's hearing ability is learned. This would apply equally to auditory discrimination and to acuity.

A study of the relationship of acuity and discrimination with a normal population is needed. Most of the work in this area has been conducted in the physics laboratories, usually with trained observers, and is not generally applicable to education.

A controlled study of the effect of the alleviation of the difficulty would be valuable.

The area approached by the articulation test is almost an untouched field in education and seems to the writer to offer seemingly limitless possibilities for needed research which would have direct bearing on education as a whole and reading in particular. So far this approach has been used almost exclusively in the testing of the acoustical properties of auditoriums and telephone circuits. The fact that persons doing this work are trained over a long period of time before they are used, suggests the possibilities in the area of teaching hearing. This area seems to offer opportunities to study the question of "how one hears" as well as "what one hears."

The combination of several of these questions gives a logical approach into another part of the field of audition which needs further study. New standards and new tests of hearing are greatly needed. A new test which will screen the high frequencies as the present group test does the low would be of tremendous value to the educator. It seems to the writer that the development of such an instrument, based on word symbols, is a logical step following further experimentation with the articulation work. It was with this in mind that the articulation technique was first employed in this study.

Whether new standards of hearing should be developed by persons interested in education or should be left to the medical field is a question. Certainly a six-year-old child should not be judged on the same level as a fifteen-year-old, any more than the fifty-year-old should be expected to have the same hearing acuity as the twenty-year-old. The fact that the medical field has unquestionably accepted, for children, the standard for normal hearing which the Bell Laboratories gave them a number of years ago, which was established on the basis of seventy-six ears, seems to the writer to give educators the right to approach the otolaryngologists and ask for their assistance in such an endeavor. Such an undertaking would need a large typical population of children of all ages. In order to carry out such a program good testing situations would need to be made available to the experimenter as would the most recent continuous frequency type of equipment. A number of the other studies mentioned above could be made in cooperation with a study of this kind but this is basic to further investigation of the relationship existing between hearing and education.

FACTORS AFFECTING EFFICIENCY OF INDUCTIVE REASONING[1]

LOUIS LONG and LIVINGSTON WELCH
College of the City of New York and the Institute for Research in Child Psychology at Hunter College

A. INTRODUCTION

In administering tests of inductive reasoning in which the subjects worked successively with three dimensional objects, pictures, and words it was observed that printed material could be used as a medium for investigating various aspects of the reasoning process in young children (2, 6). This observation led to the construction of a paper and pencil test which was administered to a group of children and which was subsequently revised and readministered on the basis of the results of the preliminary investigation. This report presents the results obtained from both of these paper and pencil tests.

The idea of assembling a series of tests graded in difficulty operated in the selection of the individual sub-tests to be included in the first battery. Consequently the test began with the simplest possible form of an inductive problem—the association between a single cause and a single effect. In successive parts the number of possible causes was increased and negative as well as positive instances of the effect were introduced. By such changes the test was gradually built up until the problems were similar to those described by Mill in his discussion of the Method of Agreement and the Joint Method of Agreement and Difference (4). Thereafter the problems involved the use of more abstract concepts.

As will be shown later some of the problems of the first pencil and paper test were too simple for any child who can read. Consequently these easier problems were eliminated and only the problems of the Joint Method type were retained. Within this framework two factors, viz., number of antecedents and level of abstractness, were varied in a systematic fashion, so that their effect on reasoning could be studied.

[1] From the Institute for Research in Child Psychology at Hunter College. This work has been aided by a grant from the American Philosophical Society. The authors are indebted to Mr. Robert Chin and to Miss Ruth Leder for their invaluable assistance with many parts of this research project.

B. COMPOSITE TEST OF INDUCTIVE REASONING

1. Description of Test

The first test will be referred to as the "Composite Test of Inductive Reasoning" since several different reasoning principles are employed in the various sub-tests. Nine different types of reasoning problems, varying widely in difficulty, were included, and there were ten problems of each type. The child's task throughout the test was to isolate the factor causing the effect in each problem. Names of animals, foods, and vehicles were used as causal factors and the presence of the effect was indicated by a plus sign (+), whereas the absence of the effect was indicated by a minus sign (—). From the directions, reproduced in Appendix I, it will be seen that a "game attitude" was created in the subject before the test began; for example, the child was told that he must imagine that a horse causes a plus. An example of each of the nine different parts of the Composite Test will be found in Table I. The correct answer has been italicized.

It will be seen that in the first sub-test an elementary type of association is involved. The effect occurs in the presence of only one antecedent. In the second part of the test the number of lines is increased to two and the effect is present in one instance and absent in another. A further increase in the number of lines occurs in the third part in which there are three lines—two negative and one positive. The first three sub-tests may be thought of as a small unit. In the fourth part double antecedents are introduced in their simplest form and there is one positive and one negative instance. In Part 5 double antecedents are again used but no antecedent occurs twice in the same line, as in Part 4. In addition the effect is present in both lines. The problems of Part 5 correspond to situations covered by Mill's Method of Agreement. In Part 6 the number of antecedents

is two, but the number of rows has been increased to three. The effect is present in two rows, but it is absent in the third. The causal factor is isolated by a process described by Mill and named the Joint Method of Agreement and Difference. Problems of Parts 4 and 5 may be thought of as leading up to those of Part 6. In Part 7 the set-up is the same as in Part 6, but a greater variety of antecedents is used from trial to trial. In the previous sub-tests only sub-human animals (horse, camel, and lion) were used as causal factors. In Parts 7, 8, and 9 species of five other classes were also used. A complete list of the words used as antecedents is presented below:

Lion, Camel, Horse, Boy, Man, Woman, Pear, Apple, Banana, Corn, Bean, Carrot, Bus, Auto, Truck, Rowboat, Sailboat, Battleship.

In the problems of Parts 6 and 7 each antecedent occurred twice, as is required by the Joint Method. It is possible, however, to have a pair of antecedents drawn from the same class instead of having identical words, thereby increasing the degree of abstractness of the reasoning situation. This was done in Parts 8 and 9. In Part 8, pairs of words were drawn from the same *first* hierarchy and in Part 9 pairs of words were drawn from the same *second* hierarchy. The words listed above form six first hierarchies (sub-human animals, humans, fruits, vegetables, land vehicles, and water vehicles) and three second hierarchies (animals, foods, and vehicles). In so far as Parts 6, 7, 8, and 9 are all of the Joint Method type they will be considered as a unit within which abstractness is increased.

In summary it may be stated that the different parts vary as to the number of antecedents, the number of rows, the number of positive and negative instances, and the level of abstractness.

2. *Procedure*

The test booklet consisted of a set of directions and nine separate pages, on each of which were ten problems and an example. The arrangement of the pages was the same for all subjects and the sequence indicated in Table I was used. The number of subjects tested at one time varied from 15 to 25. The children completed the test within a school period of one hour. Each child was given an initial test and a retest, the two periods being separated by one week.

The directions on the face of the booklet (see Appendix I) were read aloud by the experimenter and the example at the top of the sheet containing the first part of the test was placed on the blackboard in skeleton form, as in Table I. The correct answer was

TABLE I

EXAMPLES OF THE NINE TYPES OF REASONING PROBLEMS

Part 1

LION	+

Horse, Camel, *Lion*

Part 2

LION	+
CAMEL	—

Horse, *Lion*, Camel

Part 3

HORSE	+
CAMEL	—
LION	—

Lion, Camel, *Horse*

Part 4

HORSE	HORSE	+
CAMEL	CAMEL	—

Horse, Camel, Lion

Part 5

HORSE	CAMEL	+
HORSE	LION	+

Camel, Lion, *Horse*

Part 6

HORSE	LION	+
CAMEL	HORSE	—
LION	CAMEL	+

Lion, Camel, Horse

Part 7

(Same as Part 6; see test)

Part 8

BATTLESHIP	MAN	+
CORN	SAILBOAT	+
WOMAN	BEAN	—

Rowboat, Child, Carrot

Part 9

WOMAN	AUTO	+
PEAR	LION	+
BATTLESHIP	CORN	—

Bus, Bean, *Camel*

crossed out in this example. When all of the subjects had finished the problems on this page, the example for the second part was placed on the blackboard. This continued until all parts of the test were finished. The same procedure was followed when the children were retested.

3. Subjects

The Composite Test was presented to 111 subjects enrolled in the Hunter Elementary and High Schools. The Otis Group Intelligence Scale (Advanced Examination: Form A) was also administered to these children. The age range of the subjects was from 8 to 15 years. They were divided into seven age groups, one for each year between 8 and 15. The average chronological age, mental age and intelligence quotient for the seven groups are presented in Table II.

4. Results

Average Scores on the Initial Test and the Retest.—The average scores for the subjects in the various age groups indicate that as age increased there was an improvement in the performance on both the initial and the retest (Table III). The amount of improvement, as reflected by the average scores was not, however, proportional throughout the entire age range. There is a noticeable tendency for the children in the older age groups to make scores close to the maximum (90). The presence of this "ceiling effect" suggests that the test did not discriminate between the children of the upper age levels. Consequently it seemed advisable to omit some of the

TABLE II

AVERAGE CHRONOLOGICAL AGE, MENTAL AGE, AND INTELLIGENCE OF THE DIFFERENT AGE GROUPS

Age Groups	No. of Subjects	C.A. Average (Months)	M.A. Average (Months)	I.Q. Average
8	21	104	124	120
9	13	111	127	114
10	24	126	155	123
11	12	138	156	113
12	14	149	186	125
13	9	159	183	116
14	18	174	218	125
Total	111			
Average		135	162	120

TABLE III

AVERAGE SCORES ON COMPOSITE TEST

Initial Test

Parts of Tests	Age Groups 8	9	10	11	12	13	14	Average All Ages
1	6.05	7.62	9.46	9.92	10.00	9.56	10.00	8.81
2	5.67	7.69	9.58	9.33	9.64	10.00	9.78	8.67
3	7.52	7.62	9.25	9.25	10.00	10.00	9.50	8.93
4	7.81	8.00	9.79	9.67	9.93	10.00	9.72	9.22
5	6.29	7.38	9.29	9.75	9.21	9.67	10.00	8.68
6	4.95	6.31	7.83	7.75	7.50	9.11	9.94	7.50
7	5.48	6.54	8.75	8.83	9.29	9.22	10.00	8.19
8	4.29	4.77	7.21	7.92	8.43	8.66	9.28	7.05
9	3.43	4.77	6.08	6.75	7.07	6.33	7.17	5.82
All Parts	51.48	60.99	77.25	79.17	81.07	82.56	85.39	72.87

Retest

Parts of Tests	Age Groups 8	9	10	11	12	13	14	Average All Ages
1	7.62	7.77	10.00	10.00	10.00	10.00	10.00	9.29
2	8.05	8.08	9.79	9.67	10.00	10.00	10.00	9.32
3	8.29	8.08	9.67	9.83	10.00	10.00	10.00	9.36
4	8.52	8.15	9.96	9.83	10.00	10.00	10.00	9.48
5	7.00	6.85	9.54	10.00	9.79	9.67	10.00	8.91
6	5.86	7.00	9.54	9.00	9.07	9.11	10.00	8.47
7	5.90	6.77	9.96	9.08	9.07	9.44	10.00	8.58
8	4.71	4.92	8.54	6.92	8.57	9.78	9.94	7.55
9	4.10	4.85	7.21	6.67	6.64	7.22	9.00	6.50
All Parts	60.05	62.46	84.21	81.00	83.14	85.22	88.94	77.46

groups from further consideration. In an attempt to determine just which groups should be discarded and which should be retained the differences between the group means were evaluated by means of Student's "t-Test"; the resulting P values are presented in Table IV. A reliable difference was not found between the average scores of the eight and nine year groups, but there was a significant difference between the mean scores of these two groups and any other of the remaining five groups. The P values also show that the 13 and 14 year groups are not reliably different from each other. Finally, the P values obtained from inter-comparisons between the 10, 11, 12, and 13 year groups are all large enough to justify the conclusion that there is not a true difference between the average scores of these various groups. These results suggest that two homogeneous populations could be formed: one being made up of the eight and nine year groups and the other being made up of the 10, 11, 12, and 13 year groups. These populations were tested for homogeneity by using Fisher's method of analysis of variance (1, 5). It was found that in one case the variation between the subjects of the same age group was greater than the variation between the different age groups, while the reverse was true in the other case, but in neither was the preponderance large enough to lead to a rejection of the hypothesis that the two populations are homogeneous ($P > 0.05$). When these two populations were thrown together and tested for homogeneity the null hypothesis was, however, found to be untenable, since the variation between the different age groups was so much larger than the variation between children of the same age ($P < 0.01$.) In view of the finding that the 8 and 9 year groups form one homogeneous group and that the 10, 11, 12, and 13 year groups form another, it was decided to include only these groups in the discussion of the results.

The fact that two rather distinct groups emerge from an analysis of the data suggests that the tests measure abilities that are still developing at the eighth year, but which have reached their final stage of development by the tenth or eleventh year.

Analysis of the Scores on the Sub-Tests.— By considering the average scores for the different parts (sub-tests) some indication of the difficulty of the sub-tests can be obtained. Generally there is an increase in the average score followed by a decrease as the score for the last sub-test is reached. Consequently the series of tests do not appear to be graded in difficulty. The reliability of the differences between the scores of the several sub-tests was investigated by Student's "T-test." The P values will be found in Table V.

For purposes of discussion the sub-tests may be divided into three groups: (1) Parts 1, 2, and 3; (2) Parts 4 and 5; (3) Parts 6, 7, 8, and 9. The first group consisted of problems with only one antecedent, but with a varying number of lines. The second group is composed of problems with two antecedents and two lines, but with a variation in the number of positive instances in any one problem. The third group consisted of problems of the Joint Method type, but the abstractness of the material was varied.

The results obtained by comparing sub-tests within a group will be discussed first and this will be followed by a discussion of the results obtained by comparing one group of sub-tests with another. The scores on the tests of the first group were found to be not significantly different from each other except in the case of the scores of the eight year old children on Parts 2 and 3. In comparing the two sub-tests of the second group the eight year olds were found to have done significantly better on Part 4 than on Part 5, but in all other instances there was no real difference between the scores. In making comparisons between the tests in the third group

TABLE IV

P VALUES OBTAINED FROM "t-TEST" BETWEEN AVERAGE SCORES

	Year Groups					
	9	10	11	12	13	14
8 Yr. Group	>.20	<.01	<.01	<.01	<.01	<.01
9 Yr. Group		<.01	<.02	<.01	<.02	<.01
10 Yr. Group			>.60	>.30	>.20	<.02
11 Yr. Group				>.60	>.40	>.05
12 Yr. Group					>.70	>.10
13 Yr. Group						>.30

TABLE V

ANALYSIS OF SCORES ON COMPOSITE TEST—P VALUES OF DIFFERENCE BETWEEN SCORES

Parts of Test Compared.	Age Groups					
	8	9	10	11	12	13
1 vs. 2	>.40	>.80	>.30	>.20	>.20	>.30
1 vs. 3	.10	>.99	>.40	>.30	>.99	>.30
2 vs. 3	<.05*	>.80	>.10	>.90	>.20	>.99
4 vs. 5	<.02	>.20	>.20	>.90	>.10	>.30
5 vs. 6	>.05	>.05	<.05	<.05	>.05	>.30
6 vs. 7	>.30	>.50	<.05*	>.05	<.05*	>.70
7 vs. 8	>.05	>.05	>.05	>.20	<.02	>.40
7 vs. 9	<.02	<.05	<.01	<.01	<.01	<.01
8 vs. 9	>.10	>.99	>.10	>.10	>.10	<.02

Parts of Tests Compared.	Age Groups					
	8	9	10	11	12	13
1–2–3 vs. 4–5	>.10	>.90	>.60	>.40	>.10	>.80
1–2–3 vs. 6–7–8–9	<.01	<.05	<.01	<.02	<.01	>.05
4–5 vs. 6–7–8–9	<.01	<.01	<.01	<.01	<.01	<.05

* Mean of second part is reliably larger than mean of first.

Part 6 was contrasted with Part 7 and in two instances the increases in the scores of Part 7 over Part 6 were found to be reliable. The trend was present for the other age groups but the increase was not significant. The differences between Parts 7 and 8 were great enough in one instance to be considered reliable. The average scores on Part 7 were reliably higher than those of Part 9 for all groups and the scores of Part 8 were greater than those of Part 9 but the differences between the scores were unreliable except in the case of the 13 year old groups. In summary, increasing the number of lines in problems where there was only one antecedent did not make the problem more difficult; nor did the variation in the number of positive and negative instances affect performance; but increasing the level of abstractness of the Joint Method problems did reduce the efficiency.

The second type of comparison deals primarily with the differences between the three groups of sub-tests. The results of the comparisons (Table V) indicate that the first two groups contained tests of the same order of difficulty, and that these problems were definitely easier than those of the Joint Method type found in the third group. One additional comparison was made between the two separate sub-tests, Parts 5 and 6, because they represent two of Mill's methods (Method of Agreement and Joint Method of Agreement and Difference). In all instances the scores for the Method of Agreement were higher than those for the Joint Method, but in only two instances was this preponderance great enough to be considered statistically significant.

A final comment should be made about practice effect. That it operated cannot be doubted. An accurate index of the importance of this effect cannot, however, be obtained since a standard order of presentation of the sub-tests was used. A partial measure of this effect can be obtained by comparing the scores of Parts 6 and 7, since the sub-tests differ only in the variety of antecedents used. The average scores for all age groups were higher for Part 7 than for Part 6 and in two cases this difference was found to be reliable. This comparison, while inadequate, indicates that practice probably raised the scores of all of the tests, so that the scores obtained for Part 5, for example, are higher than they would have been if the subjects had started with Part 5.

Reliability of the Test.—In experimenting with tests of this type some measure of reliability must be obtained. One procedure that is often used involves testing the subjects on two different occasions and then correlating the two scores (retest reliability coefficient). In an alternate procedure the scores on half of the items of the test are correlated with the scores on the other half. Usually the scores on the odd items are correlated with the scores on the even (odd-even re-

liability coefficient). Both the retest and the odd-even reliability coefficients were obtained and they are presented in Table VI. The retest correlations for the various age groups range from 0.66 to 0.98; for the entire group (age being held constant) the correlation coefficient is 0.85. The odd-even coefficients are somewhat higher, the range being from 0.86 to 0.96; for the entire group (age constant) the correlation coefficient is 0.92. At some ages the test does not appear to be reliable if the retest correlations alone are used as an index, whereas the odd-even coefficients would suggest that the test is reliable for all of the various ages. Since it is not the purpose of the present writers to construct a standardized test, it can, we believe, be concluded that the test is reliable enough to warrant its use in an exploratory study of the reasoning ability of children.

TABLE VI

RELIABILITY COEFFICIENTS: RETEST AND ODD-EVEN

Age Groups	Retest Coefficients	Odd-Even Coefficients
8	0.76	0.93
9	0.98	0.92
10	0.66	0.94
11	0.92	0.87
12	0.84	0.86
13	0.90	0.96
8 and 9*	0.90	0.93
10, 11, 12, and 13*	0.81	0.92
All 6 Groups*	0.85	0.92

* Age held constant.

Relationship Between the Scores on the Initial Test and Chronological Age, Mental Age, and Intelligence.—The 93 subjects between 8 and 14 years of age to whom the Composite Reasoning Test was administered were also given the Advanced Examination of the Otis Group Intelligence Scale in order to relate the scores on the reasoning test to intellectual development. The correlation between I. Q. and the scores on the reasoning test was .26 (S. $E._r \pm .10$). The correlation between the Otis M. A. and the scores was .62 (S. $E._r \pm .06$); when C. A. was partialled out the correlation dropped to .31. This rather slight association between intellectual development and scores on the reasoning test can be explained partly by the fact that group tests were used. The correlations do, however, compare favorably with those obtained by the present writers in another study (3).

The combined effect of C. A. and M. A. on the scores of the reasoning test is clearly brought out by the average scores reported above. The fact that a heterogeneous group is formed when all of the subjects are included in one group, but that two homogeneous groups are formed when the 8 and 9 year old children were included in one group and the 10, 11, 12, and 13 year old children in another, undoubtedly indicates the operation of C. A. and M. A. That the relationship is not very close is shown, however, by the fact that the two homogeneous groups can be formed in spite of a variation of two years in one case and four in the other.

C. JOINT METHOD TEST OF INDUCTIVE REASONING

1. Description of Test

The results of the Composite Test clearly demonstrated that the very simple reasoning problems in the first few parts of the test would be solved by most children who were able to read. If a paper and pencil test of this type is to discriminate between children of 8 and 15 years of age it seemed necessary to concentrate on problems similar to those of Parts 6, 7, 8, and 9 of the first test (Joint Method problems). Consequently the second attempt to construct a test resulted in a new battery of 10 different parts, all of the Joint Method type. Since the form of the problems corresponds to that described by Mill in his Joint Method of Agreement and Difference, this test will be called the "Joint Method Test of Inductive Reasoning." Many of the same factors varied in the Composite Test were also varied in the Joint Method Test, but the starting point in the second was at a higher level of difficulty than in the first. Specifically, two factors, number of antecedents and level of abstractness, were varied so that their effect on reasoning could be determined.

The Joint Method Test differs from the Composite Test in several other respects. Instead of indicating the presence (or absence) of the effect by a plus (or minus) sign the word sick (or well) was used as the effect (or absence of it). The sickness was caused by various foods, including liquids, that the

hypothetical person was assumed to have eaten. A complete list of the antecedents will be found in Table VII.

The problems of Parts 1, 2, 3, and 4 are all at the lowest level of abstractness since each antecedent occurs twice. This level will be referred to as the object level. The variety of antecedents is restricted in Part 1, since they are drawn from only three classes (fowl, non-shell fish, and soft beverages); but in all other parts of the test antecedents were drawn from all of the classes listed in Table VII. The number of antecedents increased from two in Part 2, to three in Part 3, to four in Part 4. The number of lines was increased from three to four in Parts 3 and 4. This change was necessary if each antecedent was to occur twice.

TABLE VII

CLASSIFICATION OF FOODS

A. Meat		
1. Non-fowl	2. Fowl	
Beef	Chicken	Partridge
Ham	Duck	Pheasant
Lamb	Goose	Quail
Venison	Grouse	Squab
B. Fish		
3. Shell	4. Non-shell	
Clam	Bass	Perch
Crab	Carp	Salmon
Lobster	Eel	Shad
Oyster	Flounder	Trout
C. Beverage		
5. Hard	6. Soft	
Beer	Coca cola	Grapejuice
Claret	Cocoa	Milk
Port	Coffee	Soda
Sherry	Gingerale	Tea
D. Cereal		
7. Unprepared	8. Prepared	
Cream of wheat	Corn flakes	
Mush	Grapenuts	
Oatmeal	Puffed rice	
Porridge	Shredded wheat	
E. Vegetable		
9. Green	10. Non-green	
Lettuce	Beet	
Peas	Corn	
Spinach	Potato	
String beans	Squash	
F. Seasoning		
11. Spices	12. Sauces	
Cinnamon	Gravy	
Cloves	Ketchup	
Nutmeg	Tabasco	
G. Candy		
13. Hard	14. Soft	
Life savers	Bonbon	
Lollipop	Fudge	
Peanut brittle	Gum drop	
H. Fruit		
15. Citrous	16. Non-citrous	
Grapefruit	Apple	
Lemon	Peach	
Lime	Pear	
Orange	Plum	

In Parts 5, 6, and 7 no two antecedents are identical; instead two words are drawn from each of the *first* hierarchies. The number of antecedents was increased from two in Part 5, to three in Part 6, and to four in Part 7. There are three lines in Part 5 and four in Parts 6 and 7.

In Parts 8, 9, and 10 no two antecedents are identical; instead two words are drawn from each of the *second* hierarchies and the number of antecedents varies from two to four, increasing from two in Part 8, to three in Part 9, to four in Part 10.

The test was constructed so that the three following analyses could be made: (1) the effect of increasing the number of antecedents from two to four at the object level, at the first hierarchy level, and at the second hierarchy level; (2) the effect of varying the level of abstractness from the object level to the second hierarchy level when two, three, and four antecedents are used; and (3) the interaction effect of increasing the number of antecedents and varying the level of abstractness.

2. *Procedure*

Ten problems for each part of the test were typewritten on separate sheets of paper and then a test booklet was made of these 10 pages plus a page containing the directions for the test. The subjects were tested in groups of 25 to 35 children. The test was completed by the majority of the subjects within an hour.

The children were told to read the directions on the face of the test (see Appendix II) carefully before beginning the test paper. In the instructions the child was told to go on to the next test as soon as he had finished any one section. Two experimenters were always present to answer questions raised by the children.

TABLE VIII

EXAMPLES OF THE TEN TYPES OF REASONING PROBLEMS—JOINT METHOD TEST

Part 1

TROUT	EEL	WELL
CARP	TROUT	SICK
CARP	EEL	SICK

Carp, Eel, Trout

Part 2

CLAM	CRAB	WELL
CRAB	OYSTER	SICK
OYSTER	CLAM	SICK

Clam, *Oyster*, Crab

Part 3

COCA COLA	COCOA	TEA	SICK
GRAPEJUICE	GINGERALE	COCA COLA	SICK
GRAPEJUICE	GINGERALE	MILK	WELL
MILK	TEA	COCOA	WELL

Gingerale, Cocoa, *Coca Cola*, Milk, Grapejuice, Tea

Part 4

SQUAB	GROUSE	PHEASANT	CHICKEN	WELL
PARTRIDGE	QUAIL	GROUSE	PHEASANT	SICK
QUAIL	GOOSE	SQUAB	DUCK	WELL
DUCK	PARTRIDGE	CHICKEN	GOOSE	SICK

Pheasant, Quail, *Partridge*, Goose, Squab, Duck, Grouse, Chicken

Part 5

PERCH	PEAR	SICK
TABASCO	PEACH	WELL
KETCHUP	FLOUNDER	SICK

Gravy, Plum, *Carp*

Part 6

GOOSE	CLOVES	MILK	SICK
COCA COLA	CORN	GROUSE	WELL
POTATO	CLAM	GUM DROP	WELL
BONBON	CINNAMON	LOBSTER	SICK

Chicken, Fudge, *Nutmeg*, Crab, Gingerale, Beet

Part 7

HAM	PEAR	OATMEAL	SHREDDED WHEAT SICK
BEEF	LIFE SAVERS	BEER	PUFFED RICE WELL
PEACH	NUTMEG	PEANUT BRITTLE	OYSTER SICK
CRAB	SHERRY	PORRIDGE	CINNAMON WELL

Lollipop, Cloves, Clam, Claret, Venison, *Apple*, Mush, Cornflakes

Part 8

LIFE SAVERS	BEEF	SICK
BASS	QUAIL	WELL
BONBON	CRAB	SICK

Fudge, Ham, Oyster

Part 9

SHREDDED WHEAT	BEEF	CORN	SICK
CHICKEN	PLUM	NUTMEG	SICK
STRING BEANS	MUSH	FLOUNDER	WELL
OYSTER	TABASCO	GRAPE FRUIT	WELL

Quail, Ketchup, Potato, Cream of Wheat, Clam

Part 10

GROUSE	LOLLIPOP	CORN FLAKES	SHERRY	SICK
CLAM	COFFEE	PEAR	KETCHUP	WELL
SHAD	OATMEAL	LEMON	SPINACH	SICK
FUDGE	CLOVES	SQUASH	LAMB	WELL

Ham, *Puffed Rice*, Lobster, Peach, Tabasco, Gingerale, Beet, Peanut brittle.

3. Subjects

The Joint Method Test was presented to 113 children enrolled in the Hunter College Elementary School. The age range of the subjects was from 11 to 14 years. They were divided into three age groups, one for each year interval between 11 and 14. The average chronological age of the subjects will be found in Table IX.

TABLE IX

AVERAGE C. A. OF SUBJECTS TAKING JOINT METHOD TEST

Age Groups	No. of Subjects	Average CA (Months)
11 Yr.	49	136.63
12 Yr.	31	148.94
13 Yr.	33	163.55

4. Results

The average scores of the Joint Method Test are much lower than those of the Composite Test. The ceiling effect present in the results of the latter test is absent in the results of the former test, except perhaps in the case of the 13-year-old group. As in the results of the Composite Test an increase in the average score was found as age increased (Table X). The increase was not found to be significant when the 11 and 12 year groups were compared, but the increase in the score of the 13 year group over the 11 and 12 year groups was found to be reliable in both instances. The same trend, an increase in the average score as age increased, is also reflected by the sub-test scores. The differences between the sub-test scores of the 13 year group and the 11 and 12 year groups were found to be statistically significant in all instances, whereas in only three cases (Parts 6, 7, and 10) was the 12 year old group found to be definitely superior to the 11 year group.

In the discussion of the various sub-tests it was pointed out that the Joint Method Test was constructed with the idea of studying the effect of (1) increasing the level of abstractness and (2) increasing the number of antecedents on efficiency of reasoning. As mentioned above three levels of abstractness were introduced (object level, first hierarchy level, and secondary hierarchy level) and three changes in the number of antecedents (two, three, and four antecedents) were made. To

TABLE X

AVERAGE SCORES FOR JOINT METHOD TEST
Average Scores

Parts of Test	Age Groups 11	12	13
1	8.10	8.52	9.88
2	7.73	8.52	9.73
3	7.63	8.42	9.76
4	7.29	8.26	9.94
5	6.45	7.55	9.42
6	5.76	7.13	8.30
7	4.88	6.58	8.61
8	6.94	7.55	9.52
9	4.76	5.74	8.30
10	3.98	5.26	8.09
All Parts	63.51	73.52	91.55

obtain an index of the influence of level of abstractness the scores from the different levels were paired with each other and Student's "t-Test" was used to test the significance of the difference between them. The comparison always involved problems with the same number of antecedents so that this particular variable is ruled out in the present analysis. The comparisons made are listed in Table XI along with the P values.[2] That increasing the abstractness of the material affected performance can hardly be doubted since in 16 out of 18 instances the average scores decreased as the abstractness of the test increased and in 15 of these the differences between the average scores are significant. The three cases where the differences are not reliable occur when only two antecedents are used.

The effect of increasing the number of antecedents was studied by regrouping the data so that the abstractness was held constant. For example, the average score for the test with two antecedents was compared with that for the test with three antecedents at the object level. The values of P for the 18 different comparisons are presented in Table XII. In nine of the comparisons a reliable difference was found between the scores, and the lower average was always associated with the sub-test having the greater number of antecedents.

A final comparison is concerned with the combined effect of the two factors of increasing levels of abstractness and increasing number of antecedents. In the two separate analyses the effect of increasing the number

[2] The data from the 13 year old group were not analyzed in this fashion since the differences between average sub-test scores are so small and furthermore the average scores are very near the maximum score of 10.

TABLE XI

RELIABILITY OF DIFFERENCE BETWEEN SCORES ON SUBTESTS HAVING SAME NUMBER OF ANTECEDENTS

COMPARISON:	P-Values Age Groups 11	12
Part 2 vs. Part 5	<.01	<.02
Part 2 vs. Part 8	<.05	>.05
Part 5 vs. Part 8	>.05	>.99
Part 3 vs. Part 6	<.01	<.01
Part 3 vs. Part 9	<.01	<.01
Part 6 vs. Part 9	<.01	<.01
Part 4 vs. Part 7	<.01	<.01
Part 4 vs. Part 10	<.01	<.01
Part 7 vs. Part 10	<.01	<.01

of antecedents was studied at the object level, then the first hierarchy level, and finally the second hierarchy level. Instead of making three separate comparisons the results from all three levels were thrown together and the reliability of the difference between sub-tests with two, three, and four antecedents was calculated. The average score decreased as the number of antecedents increased and this decrease was found significant in all cases except for the 12 year group when the number of antecedents was changed from three to four. When the same type of analysis was made for hierarchy level the results indicate that the scores decreased as the level of abstractness increased and this decrease was found to be statistically reliable in all instances.

In an attempt to evaluate the relative amount of increase in difficulty caused by varying the level of abstractness and the number of antecedents the total variance for the 11 and 12 year groups was separated into that produced by subjects, antecedents, abstractness, and the interaction of these factors. When the F ratio (5), used to test the significance of the difference in variance, was made by comparing the variance of antecedents or abstractness with the triple interaction variance a P value of less than 0.01 was obtained. The interpretation that both variables produced more than a chance variation can be offered, but the differentiation between antecedents and abstractness was not possible. Therefore an F ratio was made by comparing these two variables directly. In both the 11 and 12 year groups the amount of variation caused by abstractness was greater than that caused by antecedents, but in neither case was this preponderance large enough to meet the criterion of being statistically significant.

TABLE XII

RELIABILITY OF DIFFERENCE BETWEEN SCORES ON SUBTESTS AT SAME LEVEL OF ABSTRACTNESS

Comparison:	P-Values Age Groups 11	12
Part 2 vs. Part 3	>.50	>.70
Part 2 vs. Part 4	>.05	>.20
Part 3 vs. Part 4	>.05	>.50
Part 5 vs. Part 6	<.05	>.30
Part 5 vs. Part 7	<.01	<.05
Part 6 vs. Part 7	<.01	>.10
Part 8 vs. Part 9	<.01	<.01
Part 8 vs. Part 10	<.01	<.01
Part 9 vs. Part 10	<.01	>.05

D. DISCUSSION OF THE RESULTS

The two paper and pencil tests represent attempts to construct a series of inductive problems graded in difficulty. The plan of calibration was purely logical and involved changes of a qualitative as well as a quantitative nature. Examples of the latter type of change are increases in the number of antecedents, variations in the number of rows, and changes in the proportion of positive and negative instances in the problems; examples of the former are shifts from concepts of the object level to those of the first or second hierarchies. In the first six parts of the Composite Test several changes of the quantitative type were made. Performance on the whole was not greatly affected. The average scores for any two successive parts of the test were seldom found to be reliably different. The results suggest that the children were functioning so efficiently that such quantitative changes as were introduced were not drastic enough to upset their performance. The introduction of qualitative changes in Parts 8 and 9 did, however, cause performance to drop rather sharply—but the problems of these parts are both quantitatively and qualitatively different from those of Parts 1 and 2. Consequently the decrease in the scores can most appropriately be attributed to the combination of these two types of change.

Essentially both types of change were introduced in the Joint Method Test, but the level at which they operated was initially higher. The number of antecedents and rows

was varied but the starting point was two antecedents and three rows instead of one antecedent and one row. Shifts were also made from concepts of the object level to those of the first and second hierarchies. As in the Composite Test the qualitative changes seemed to affect performance more than the quantitative, but the effect produced by quantitative changes was greater in the Joint Method Test than in the Composite Test. This result might be interpreted to mean that at a relatively simple level of performance, modifications of the problem situation in quantitative terms, and within certain limits, do not affect the ease of solution; whereas, at more difficult levels of performance, accuracy may be decreased merely by increasing the number of elements involved even though the basic method for arriving at the solution remains unchanged. It might be said, therefore, that whether or not a quantitative change will affect performance depends upon the original difficulty of the problem on which such a change is superimposed.

In the Joint Method Test the qualitative and quantitative factors operated separately as well as jointly so that it was possible to determine the *relative* effectiveness of quantitative and qualitative changes. As might be anticipated, the greatest decrease in the scores occurred when the combined influence of these two factors was present. In other words, when the number of antecedents was increased at the same time that a higher level of abstractness was introduced, performance was more adversely affected than when either one of these changes was introduced by itself. (This finding is graphically presented for the 11 year group in Fig. 1.) When the decrease in the scores which was attributable to the operation of the qualitative factor was compared with that resulting from the operation of the quantitative factor, the former was found to be greater, but the difference was not statistically reliable.

A final point concerns the question of whether reasoning at a higher level of abstractness is always more difficult than at a lower level. There may actually be instances in which working at the second hierarchy level is easier than working at the first hierarchy level. Whether or not this is true may depend, for example, on the construction of the hierarchies. Since a hierarchy can be started from any point it may in a given instance happen that some of the categories at higher levels are more familiar than those at the lower levels. This condition could conceivably be effective in determining whether working at the first hierarchy level is more difficult than working at the second. In the Composite Test the first hierarchy concepts are used more often than are the second hierarchy concepts. For example, boy, man, and woman are more frequently thought of as humans than as animals. In the Joint Method Test the conditions are reversed—the second hierarchy concepts are used more frequently than are the first. For example, beef and ham are thought of as meats before they are thought of as nonfowls. Chicken and ham are more apt to be included in the same category than are man and horse, or so it seems to us. At any rate, this explanation is offered to cover the following observation: in the Joint Method Test, performance on the part of the test dealing with second hierarchy concepts was occasionally better than that on the part dealing with first hierarchy concepts. The average scores on the Joint Method Test reflect a tendency for the children to react in the same fashion, since the scores for the second hierarchy test with two antecedents are higher than (or as high as) the scores for the first hierarchy test with two antecedents. These differences, however, are not reliable as was pointed out above.

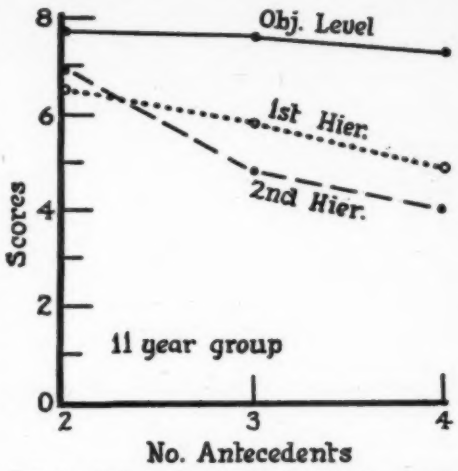


FIG. 1—Average scores for the 11 year group on the Joint Method Test.

E. Summary and Conclusions

Two paper and pencil tests of inductive reasoning were administered to children varying in age from 8 to 15 years. The Composite Test of Inductive Reasoning was given to 111 children and the Joint Method Test of Inductive Reasoning was given to another group of 113 children.

The two tests were graded in difficulty. The Composite Test involved a variety of inductive problems ranging from an elementary type of association to situations covered by Mill's Joint Method of Agreement and Difference. The Joint Method Test began with the Joint Method situation.

The difficulty was increased by making changes of a quantitative and qualitative type. Quantitative changes refer to increases in the number of antecedents or in the number of rows, and to a variation of the proportion of positive and negative instances in a problem. Qualitative changes refer to shifts in the abstractness of the concepts that were employed in the reasoning situations. In some instances the concepts were at the object level, whereas in others the concepts were at the first or second hierarchy level.

The quantitative changes introduced in the Composite Test were so elementary that performance was not noticeably affected. When the qualitative changes were superimposed upon the quantitative, a definite drop in the scores occurred.

The results of the Joint Method Test show clearly that both types of change will cause a decrease in score. The decrement produced by increasing the abstractness of the concepts was greater than that produced by increasing the number of antecedents, but the difference was not great enough to meet the criterion of being statistically reliable.

References

1. Fisher, R. A. *Statistical Methods for Research Workers*. London: Oliver & Boyd, 1934. Pp. xii + 319.
2. Long, L., and Welch, L. "Reasoning Ability in Young Children. *Journal of Psychology*, XII (1941), 21–44.
3. Long, L., and Welch, L. "Influence of Levels of Abstractness on Reasoning Ability," *Journal of Psychology*, XIII (1942), 41–59.
4. Mill, J. S. *A System of Logic*. London: 1843. Vol. I, p. 458.
5. Snedecor, G. W. *Statistical Methods*. Ames, Iowa: Collegiate Press, 1938. Pp. xiii + 388.
6. Welch, L., and Long, L. "Comparison of the Reasoning Ability of Two Age Groups," *Journal of Genetic Psychology* (in press.)

APPENDIX I

Reasoning Game

Do children like to go on errands? Most of us enjoy doing things for other people. Let's imagine that your teacher sends you to a new room in the school building to get something—say a pencil. Now as you go into the room you see three light switches on the wall. You turn on the one nearest the door and the light goes on. Then you get the pencil and return to your class room. As you give the pencil to the teacher, she asks the following question: "Which of the three switches made the light go on?" You would say the one nearest the door, wouldn't you? Now we want you to find out what makes a plus (+) rather than what makes a light go on. If you see a plus (+) then that means something happened. If you see a minus (—) that means it did not happen. Sometimes you will see that a horse gives a plus (+). Then you will be asked which one of the following will make a plus (+): lion, horse, camel. In this case the horse will make a plus (+). In playing this game you must imagine that the animal alongside of a plus (+) can make the plus (+). Then we want you to select from a group of three animals the one which you think is making the plus (+). You can find this out by thinking—don't guess.

There are several different pages, but do not turn over to these until you receive the signal. When you finish a page put your pencil down and wait for the signal to turn to the next page. An example for each new page will be put on the blackboard.

This is not a test to see how fast you can work. You will be given plenty of time—work carefully.

Now write your name after the word "Name."

APPENDIX II

Test in Causal Reasoning

On the basis of the information given in the following situations you are required to discover what particular kind of food or drink is the cause of sickness in each instance. *Do not attempt to rely on your experience in solving these problems.* For instance, do not decide that candy in one problem is the cause of sickness because candy made *you* sick. You will fail the test if you work on this basis. In order to solve the problems you must apply methods of logical reasoning. An example correctly solved, is presented at the top of each page. The example on the first page is presented here so that instructions as to how you are to read the problems can be given:

TROUT	EEL	WELL
CARP	TROUT	SICK
CARP	EEL	SICK

Carp, Eel, Trout.

This is an abbreviated way of saying that the person ate a *Trout* and an *Eel* and was *Well;* on another day he ate a *Carp* and a *Trout* and was *Sick;* on another day he ate a *Carp* and an *Eel* and was *Sick*. Which one (The Carp, Eel, or Trout) made him sick? Notice that in any one problem you are to read *across* the page, *not down* the page.

A brief definition of the various foods (or drinks) will be found in the last page of this test. Consult this sheet if you are not familiar with a term.

This is *not* a test to see how fast you can work. You will be given plenty of time. As you finish one page go right on to the next. Work carefully.

You may cross out or mark up any item on the test sheets you desire, if it aids you in any way.

FACTORS CONTRIBUTING TO ACHIEVEMENT IN THE STUDY OF FIRST SEMESTER COLLEGE GERMAN[1]

R. P. LARSEN, J. R. WITTENBORN and E. G. CIESECKE
Personnel Bureau, University of Illinois

Many investigations have been reported regarding the prediction of success in modern language. Algernon Coleman and Clara Breslove King (1) have stated, however, that there is less agreement than might be expected in regard to the importance of various factors contributing to language achievement. The present study was undertaken to investigate the importance of factors commonly believed to be pertinent to success in the study of first semester German at the University of Illinois.

The criterion for success was the total score of the *Cooperative German Test, Elementary Form O.* This test was administered during the first week of the second semester, 1939–40, to 130 Liberal Arts and Sciences freshmen who had completed the first semester and who were enrolled in the second semester of elementary German.

In the first semester German course, Part I of the textbook, *Ich lerne Deutsch* (4) was covered. The teaching procedure based on this text may be described as a modified reading method. The primary goal of the course is to develop the student's ability to read German with ease and pleasure.

The test data included scores made on the Freshman Guidance Examination Battery for freshmen entering the College of Liberal Arts and Sciences and their scores made on other selected tests which were administered at the beginning of the study.

The Freshman Guidance Examinations given in the fall of 1939 included the following: *English Training, Iowa Placement Examinations; Foreign Language Aptitude, Iowa Placement Examinations; Iowa High School Content Examination; Iowa Silent Reading Tests, New Edition.*

The additional tests which were given during the first week of the second semester were: *Word-Number* (5), *Initials Recall* (5), *Auditory Word-Number,*[2] *Auditory Initials Recall,*[2] *Backward Writing* (5), *Scattered X's* (5), *Number Patterns* (5), *Syllogism,*[3] *Alphabet* (3), *Number Three,*[4] *Number Five,*[4] and the *Otis Quick-Scoring Mental Ability Test, Gamma form.*

The *Word-Number, Initials Recall, Auditory Word-Number,* and *Auditory Initials Recall* tests measure rote memory, the *Backward Writing* and *Scattered X's* are visual perception tests, the *Number Patterns* and *Syllogism* tests are considered to involve reasoning, the *Alphabet, Number Three,* and *Number Five* were constructed as measurements of attention, and the *Otis Quick-Scoring Mental Ability Test* is a scholastic aptitude test.

In addition to the test battery, each subject was given a standard interview. This interview included items concerning high school background, present goals and interests, use of time, and general methods of study. A prepared form was used by the interviewer in conducting and recording the interview. This form served as a convenient record and also made it possible for the interviewer to question each student in a uniform order. All interviews were conducted by the same person.

Immediately preceding each interview, each student filled out a questionnaire containing personal and scholastic items.

A method of matched groups was employed to eliminate the effect of general scholastic aptitude as measured by the *Otis Quick-Scoring Mental Ability Test,* which was found to correlate .368 with the criterion in the original population of 85 students. This was

[1] Since the work reported here was carried out at the Personnel Bureau, the writers take this opportunity of expressing their indebtedness to the Personnel Bureau staff for their advice and generous assistance.

[2] Prepared specifically for this study. The tests were similar to Thurstone's *Word-Number and Initials Recall* tests except that the presentation was auditory rather than visual. The time limits for the auditory presentation were the same as those for the visual presentation.

[3] A true–false test consisting of a series of syllogistic statements, some with logical conclusions, others with illogical conclusions.

[4] Lists of letter and number combinations are presented by means of phonograph records. Certain rules govern the response to these combinations. The subject must promptly apply the rule applying to the combination heard. More information regarding these tests may be secured from J. R. Wittenborn.

done by distributing the Otis scores into groups each with a range of five raw score points. The students in each Otis score group having the highest and lowest German achievement scores were selected. Those students, 27 in number, having the high German scores constituted the High Achievement Group; those having the low scores, 27 students, made up the Low Achievement Group. Thus, the students in each group were matched on Otis scores, and were so discrepant in German scores that the lowest score for the High Achievement Group was sixteen points higher than the highest score in the Low Achievement Group.

Since participation in the program was voluntary, it was possible to secure complete test data from only 85 Liberal Arts and Science freshmen. Six of these students could not be included because they had had previous experience with German. Since there were 27 students in each of the matched groups, this study is concerned with 54 students, or 68 per cent of the population for whom complete data were available.

Test Results

From the tabular presentation of the German achievement and Otis test data in Table I, it is apparent that it has been possible to select two groups of students almost identical in scholastic aptitude as measured by the Otis test (C.R. = .10), but widely discrepant in German achievement (C.R. = 13.18). From the study of such selected cases, knowledge of a great portion of the students can be only inferential; on the other hand, such a procedure can be expected to throw into sharp relief the abilities and habits peculiar to the high and the low achieving students.

The performance of high and low achievers on the Freshman Guidance Examinations is shown in Table II. Reading performance is shown in terms of rate and comprehension, and the total score for the *High School Content Test* is given. The total and sub-test scores for the *English Training* and the *Foreign Language Aptitude Tests* are given.

These data strongly suggest that the *Foreign Language Aptitude Test* is of limited value for the prediction of achievement in the first semester of German; the difference between the means of the high and low achievement groups was 14.59. The critical ratio of this difference is 2.29, which does not meet the criterion for a statistically significant difference. Of equal interest is the greater predictive value of the *English Training Test* for achievement in German. The difference between the mean scores for the high and low achievement groups is 33.22; the critical ratio of this difference is 4.01.

From the data in Table III, it appears that the tests assumed to measure memory, reasoning, concentration or attention, and perception are of no value in differentiating between high and low achievers in first semester college German.

Discussion of Test Results

From the tabulation of the test data in Table III, it is apparent that memory, reasoning, attention, and perception test scores for the high achievement group are not significantly higher than those test scores for the low achievement group. The writers are

Table I

Relative Performance of the High and Low Achievement Groups on the Cooperative Elementary German Test and on the Otis Quick-Scoring Mental Ability Test[a]

	Mh*	Ml*	Mh-Ml	S. D. Diff.	C. R.
A. Cooperative German Test					
Reading	26.59	15.25	11.34	1.30	8.72
Vocabulary	32.07	19.96	12.11	1.07	11.32
Grammar	23.59	15.18	8.41	.87	9.67
Total Score	82.66	50.50	32.16	2.44	13.18
B. Otis Quick Scoring					
Total Score	59.15	58.90	.25	2.42	.10

* Mh and Ml are each based on 27 cases.

[a] In Tables I, II, and III, Mh and Ml stand for the mean score of the high and low achievement groups respectively. Mh-Ml, S.D.Diff., and C.R. stand for the difference between the means, the standard deviation of this difference, and the critical ratio of the difference.

TABLE II

RELATIVE PERFORMANCE OF THE HIGH AND LOW ACHIEVEMENT GROUPS ON SCORES FOR THE FRESHMAN GUIDANCE EXAMINATION BATTERY

	Mh*	Ml*	Mh-Ml	S. D. Diff.	C.R.
A. English Training Test					
Spelling	35.26	26.52	8.74	2.90	3.01
Punctuation	35.70	27.92	7.78	2.66	2.93
Sentence Usage	31.63	18.15	13.48	2.93	4.60
Word Connotation	37.56	33.33	4.23	2.07	2.04
Total Score	140.15	106.92	33.22	8.29	4.01
B. Foreign Language Aptitude					
Grammar	31.55	24.41	7.14	2.73	2.61
Reading from Context	23.74	22.19	1.55	1.40	1.11
Esperanto Grammar	20.70	19.00	1.70	1.47	1.16
Translation	16.52	12.18	4.33	2.42	1.79
Total Score	92.52	77.93	14.59	6.35	2.29
C. High School Content					
Total Score	157.81	136.88	20.93	7.94	2.64
D. Iowa Silent Reading					
Rate	36.89	32.96	3.93	1.86	2.11
Comprehension	185.89	169.96	15.93	7.91	2.01

*Mh and Ml are each based on 27 cases.

TABLE III

RELATIVE PERFORMANCE OF THE HIGH AND LOW ACHIEVEMENT GROUPS ON MISCELLANEOUS ABILITY TESTS

	Mh*	Ml*	Mh-Ml	S. D. Diff.	C.R.
Auditory Word-Number	14.52	14.74	—.22	1.23	—.18
Auditory Initials Recall	13.52	11.85	1.67	1.42	1.17
Word-Number	7.30	9.33	—2.03	1.32	—1.54
Initials Recall	12.04	11.96	.08	1.75	.04
Number Patterns	9.55	9.10	.46	1.11	.41
Syllogism	13.63	14.81	—1.18	1.06	—1.11
Alphabet	25.96	25.30	.66	.201	.33
Number Five	68.22	64.04	4.18	4.63	.90
Number Three	77.30	72.18	5.11	5.79	.85
Backward Writing	36.15	34.77	1.38	2.25	.61
Scattered X's	21.15	22.74	—1.59	1.64	— .97

*Mh and Ml are each based on 27 cases.

not prepared to say that the abilities of memory, reasoning, attention, and perception as measured by the tests used, play no part in acquiring the knowledge of the German langauge as measured by the *Cooperative German Test.* It is possible, however, that the amount of these abilities necessary for the acquisition of any degree of proficiency in college German as measured by the achievement test was possessed by all of the subjects, or it is possible to assume that the tests of these abilities were either too easy or too difficult and as a consequence failed to measure the ability range crucial to performance on the *Cooperative German Test.*

That reading rate and comprehension play some part in German achievement are suggested by the critical ratios of 2.11 and 2.01, respectively. The size of these critical ratios does not meet the criterion for statistical significance; however, it does indicate that the high achievement group is much more likely to score higher on the reading test than the low achievement group.

The total score on the *High School Content Test* is substantially higher for the high achievement group as indicated by the critical ratio of 2.64.

The *Foreign Language Aptitude Test,* consisting of three sub-tests involving artificial language and Esperanto and one sub-test in-

volving English grammar, failed to discriminate significantly between the high achievement and the low achievement groups. It is to be further noted that the only sub-test of the *Foreign Language Aptitude Test* in which the mean score of the high achievement group approached being significantly larger than the mean score for the low achievement group was the sub-test involving basic English grammar. This is consistent with Kamman's (2) findings to the effect that the grammar sub-test correlated most highly with a four semester German grade average. On the basis of these data it appears that the *Foreign Language Aptitude Test* leaves much to be desired as a predictor of achievement in first semester German.

The relative performance of the high and the low achievement groups on the *English Training Test* reveals a need for further study. The mean scores of the high achievement group on all but the fourth of the *English Training* sub-tests are significantly higher than the respective mean scores made by the low achievement group. The total score difference between the means of the high and the low achievement groups divided by the standard deviation of this difference gives a critical ratio of 4.01. A significant question arises here concerning the importance of knowledge of English grammar in learning German. Did those students who were able to make a high score on the *English Training Test* also make higher scores on the German test because of some basic ability common to both English and German achievement, or did they make high scores on the German test because they had been better trained in English?

Interview and Questionnaire Results

It is recognized that the reliability of the interview and questionnaire is dependent upon the accuracy of the students' reports. Every precaution was taken, however, to get as careful and as reliable a report as possible from the student.

The tabular presentation of the data for the questionnaire and interview material is in terms of the per cent of each group possessing the habit or following the method involved in each inventory or interview item. In some instances, the answers had a numerical value; such cases are listed separately, and mean values for each group are used as in the tabulation of the test results. The tabulation presents the per cent or proportion of affirmative responses given by the high achievement group, the proportion of affirmative responses given by the low achievement group, the difference between the proportions for the high and low groups, the standard deviation of this difference and the critical ratio.

For all differences, critical ratios are expressed; a difference is said to be statistically significant when the critical ratio is as large as 3.00. Many differences do not meet this criterion of significance, however, in general, greater significance can be attributed to the difference giving the larger critical ratios. The relative significance of the differences between the high and low achievement groups will be considered.

The results of the interview and questionnaire material are presented in Tables IV, V, VI, and VII, which are the respective tables for items classified as bearing on high school background, goals, interests, use of time and methods of study.

The data in Table IV bearing on high school background indicates that the amount of general high school study and extra reading, and the number of years of language study in high school are similar for the two achievement groups. This observation, that the amount of general high school study and extra reading, and the number of years of language study in high school do not materially affect achievement in college German, is not in keeping with opinion; but, quite according to expectation is the fact that in this study the per cent of low achieving students reporting that high school language study aided them in the study of college German is significantly less, (C.R. = 2.95) than the per cent of the high achievement group making this report.

The items concerning the question of the importance of goals and attitudes are of two types, those which are related to the student's general outlook and those which are specific to German. It is seen from Table V that definite reasons for attending college, steadfastness of professional goals, and grade-point expectancy bear little or no relationship to scores made on the German achievement test. On the other hand, that significantly more of the high achievement group were interested in German, desired to master

TABLE IV

RELATIVE STANDINGS OF THE HIGH AND LOW ACHIEVEMENT GROUPS ON ITEMS RELATING TO HIGH SCHOOL BACKGROUND[a]

	%h*	%l*	%h-%l	S. D. Diff.	C. R.
A. Did you study in high school?	.19	.22	—.08	.14	—.21
B. Did you do extra reading in high school?	.33	.41	—.08	.13	—.61
C. Did high school language study help you?	.74	.37	.37	.13	2.95
	(Mh)	(Ml)	(Mh-Ml)		
D. Years in high school language	2.43	2.52	—.09	.28	—.34
E. High school grades reported	4.44	4.11	.34	.17	1.94

* h and l are each based on 27 cases.

[a] In Table IV, as in Tables V and VI, %h and %l stand for the per cent of the high or low achievement group possessing a given habit or answering a given item in the affirmative. %h-%l stands for the difference between the proportion, S. D. Diff. stands for the standard deviation of the difference, and the C. R. stands for the critical ratio of the difference. The column headings (Mh), (Ml), and (Mh-Ml) show the mean standing of the high and low groups, and the differences between the mean standing of the high and low groups for questions that have responses carrying a numerical value.

TABLE V

RELATIVE STANDING OF THE HIGH AND LOW ACHIEVEMENT GROUPS ON ITEMS RELATING TO GOALS AND INTERESTS

	Mh*	Ml*	Mh-%l	S. D. Diff.	C. R.
A. Did you have a definite reason for college?	.63	.52	.11	.13	.82
B. Was German pertinent to your goals?	.78	.48	.30	.12	2.40
C. Did you aim to master the language?	.63	.22	.41	.12	3.34
D. Did you view assignments as steps to mastery?	.81	.33	.48	.12	4.09
E. Have your goals changed while in college?	.85	.78	.07	.11	.66
F. Did you aim to master each lesson?	.78	.52	.26	.12	2.08
G. Were you interested in German?	.96	.41	.55	.10	5.42
H. Did you view assignments as isolated units?	.52	.74	—.22	.13	—1.72
	(Mh)	(Ml)	(Mh-Ml)		
I. Average college grades expected	4.00	3.75	.25	.15	1.65

* h and l are each based on 27 cases.

the language, and viewed each assignment as a step toward mastery, is indicated by the respective critical ratios of 5.42, 3.34 and 4.09. More low achieving students were inclined to view assignments as isolated units; however, it should be kept in mind that this difference is a small one.

Of the numerous items in Table VI relating to the use of time, only two show differences between the high and low achievement groups that approach statistical significance: item E, covering the lesson in the time set aside, critical ratio 2.98; and item F, making daily preparation, critical ratio 2.75. From item K it is seen that the members of the low group tend to spend more time on German preparation than do the members of the high group. The difference between the mean times for the two groups is approximately one and one-half hours per week; the standard deviation of this difference is relatively large and the critical ratio is far from being statistically significant.

It is revealed from Table VII that significantly more of the high achievement group studied corrections on papers returned, and read ideas rather than words in translating German. Also, the difference between the per cent of the high achievement group studying vocabulary from context approaches significance as indicated by the critical ratio of 2.36. Of additional interest is the slight tendency for the low achievement group to follow the methods described in items G, H, I, and K.

SUMMARY AND CONCLUSIONS

Freshmen in the College of Liberal Arts and Sciences who had just completed one semester in German were given the elementary *Cooperative German Test*, the *Otis Quick-Scoring Mental Ability Test*, and a

TABLE VI

RELATIVE STANDING OF THE HIGH AND LOW ACHIEVEMENT GROUPS ON ITEMS RELATING TO THE USE OF TIME

	%h*	%l*	%h-%l	S. D. Diff.	C. R.
A. Did you stick to a time study budget?	.33	.22	.11	.12	.91
B. Did you spend your study time in one stretch?	.78	.78	.00	.11	.00
C. Did you warm up for study easily?	.78	.67	.11	.12	.91
D. Did you attend class regularly?	1.00	1.00	.00	.00	.00
E. Did you cover the lessons in the time set aside?	.96	.67	.29	.10	2.98
F. Did you make daily preparation?	1.00	.78	.22	.08	2.75
G. Did you use frequent review?	.30	.19	.11	.12	.95
H. Were you prompt with written work?	.96	.96	.00	.05	.00
	(Mh)	(Ml)	(Mh-Ml)		
I. Number of hours worked per week for pay	7.41	9.33	—1.93	3.48	—.55
J. Number of hours spent per week in activities	3.74	5.00	—1.26	1.33	—.95
K. Number of hours set aside for German per week	7.96	9.22	—1.26	.85	—1.47
L. Minutes of daily oral reading	41.48	40.00	1.48	15.01	.10

* h and l are each based on 27 cases.

TABLE VII

RELATIVE STANDING OF THE HIGH AND LOW ACHIEVEMENT GROUPS ON ITEMS RELATING TO GENERAL STUDY METHODS

	%h*	%l*	%h-%l	S. D. Diff.	C. R.
A. Did you easily adapt to study in college?	.33	.37	—.04	.13	—.31
B. Was your study environment good?	.70	.70	.00	.12	.00
C. Was it easy for you to concentrate?	.52	.48	.04	.14	.29
D. Did you cooperate in class?	.93	.93	.00	.07	.00
E. Did you find pronunciation quite difficult?	.52	.44	.08	.14	.59
F. Did you drill paradigms silently?	.52	.52	.00	.14	.00
G. Did you drill orally?	.22	.33	—.11	.12	—.91
H. Did you drill paradigms by writing?	.48	.67	—.19	.13	—1.44
I. Did you study vocabulary by silent reading?	.37	.56	—.19	.13	—1.42
J. Did you study vocabulary by oral drill?	.22	.19	.03	.11	.26
K. Did you study vocabulary by writing?	.22	.41	—.19	.12	—1.54
L. Did you study vocabulary in context?	.56	.26	.30	.13	2.36
M. Did you keep a class notebook?	.26	.37	—.11	.13	—.88
N. Did you study corrections?	.89	.52	.37	.11	3.26
O. Did you learn the why of drill or exercises?	.89	.74	.15	.10	1.44
P. Did you associate new words with familiar?	.74	.63	.11	.13	.88
Q. Did you read for ideas rather than words?	.89	.41	.48	.11	4.28

* h and l are each based on 27 cases.

number of miscellaneous tests assumed to be measures of memory, reasoning, attention, and perception.

From the 79 students for whom complete data were available, two groups of 27 students each were selected. These two groups were matched in scholastic aptitude, but were so discrepant in German achievement that the lowest score of the high achievement group was sixteen points higher than the highest score of the low achievement group.

The students, further, filled out a questionnaire and met with a counselor in a standard interview for the purpose of discussing the high school background, goals and interests, use of time, and general methods of study.

This investigation is concerned with a comparison of the relative standing of the high and low achievement groups on all of the measures employed, including the Freshman Guidance Examinations.

The following statements are a result of a comparison of the test performance of the high achievement group with the low achievement group:

1. When scholastic aptitude, as measured by the *Otis Quick-Scoring Mental Ability Test*, is held constant, students having high German achievement scores tend to have significantly higher scores on the *English Training Test* than students having low German achievement scores.
2. When scholastic aptitude is held constant, students having high German achievement scores do not have *Foreign Language Aptitude* scores significantly higher than students having low German achievement scores. This is also true for the *High School Content Examination* and for the rate and comprehension scores of the *Iowa Silent Reading Tests*.
3. Scores on the tests employed in this study as measures of memory, reasoning, attention and perception revealed no relationship to either high or low achievement in German.

When the high and low achievement groups were compared on the basis of responses to interview and questionnaire items, the following results were obtained:

1. Significantly more of the high achievement group reported that high school language study helped them in their present study of German.
2. The items relative to goals and interests revealed that significantly more of the high achievement group wanted to master the German language, viewed each assignment as a step toward mastery, and were interested in German.
3. In regard to the use of time, it was found that the high achievement group was more likely to cover their lessons in the time set aside and make daily preparation than the low achievement group.
4. In the area of general study methods, it was found that significantly more of the high achievement group studied corrections and significantly more read for ideas rather than for words. There was also a definite tendency for the high achievement group to study their vocabulary in context.

References

1. Coleman, A., and King, Clara B. *Encyclopedia of Educational Research*. New York: the Macmillan Company, 1941. Pp. 522.
2. Kamman, J. F. *A Statistical Analysis of the Foreign Language Aptitude Test*, Form B of the Iowa Placement Examination. Unpublished Masters Thesis, State University of Iowa, 1939.
3. Phillip, B. R. "The Measurement of Attention," *Studies in Psychology and Psychiatry*, II (1928).
4. Ross, Flora, and Ross, Ada E., and Aron, A. W. *Ich Lerne Deutsch*. New York: Harper, 1935.
5. Thurstone, L. L. *Tests for Primary Mental Abilities*, Experimental Edition, Washington, D. C.: American Council on Education, 1938.

Ende.